

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:25:31 ; Search time 20 Seconds  
(without alignments)  
801.775 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDAPTMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	285	9.8	625	4	US-09-590-020-7
2	272	9.4	623	4	US-09-590-020-4
3	269	9.3	623	4	US-09-590-020-2
4	262.5	9.1	622	4	US-09-590-020-6
5	149	5.1	2206	1	US-07-852-260-2
6	149	5.1	2206	2	US-08-461-503-2
7	149	5.1	2206	4	US-08-465-250-2
8	106	3.7	1302	4	US-09-423-890-2
9	106	3.7	1493	4	US-09-423-890-8
10	106	3.7	1593	4	US-08-628-829-4
11	106	3.7	2318	4	US-09-091-219-24
12	103.5	3.6	1170	4	US-09-749-588-2
13	102.5	3.5	588	1	US-08-460-860-4
14	102	3.5	1048	4	US-09-171-699-10
15	100	3.5	544	4	US-09-615-192A-349
16	98	3.4	1209	4	US-09-749-588-4
17	97	3.3	956	4	US-09-134-078-63
18	97	3.3	3443	2	US-08-416-603-2
19	95.5	3.3	202	2	US-08-416-603-10
20	95	3.3	2232	4	US-09-091-219-25
21	95	3.3	2247	4	US-09-091-219-2
22	94.5	3.3	2004	1	US-08-375-709-15
23	94.5	3.3	2004	1	US-08-752-929-15
24	94.5	3.3	2004	4	US-09-090-793-9
25	94	3.2	1035	4	US-09-206-942-73
26	94	3.2	1477	1	US-08-038-682-4
27	94	3.2	1477	1	US-08-302-832-4

28	34	3.2	1477	2	US-08-530-198-4	Sequence 4, Appli
29	34	3.2	1477	2	US-08-469-880-4	Sequence 4, Appli
30	34	3.2	1477	2	US-08-728-470-4	Sequence 4, Appli
31	34	3.2	1477	2	US-08-617-697-4	Sequence 4, Appli
32	34	3.2	1477	4	US-08-719-641-4	Sequence 4, Appli
33	34	3.2	1477	4	US-09-206-942-71	Sequence 71, Appli
34	93.5	3.2	907	3	US-08-783-774-2	Sequence 2, Appli
35	93.5	3.2	907	4	US-09-328-599A-1	Sequence 1, Appli
36	93.5	3.2	907	5	PCT-US95-04611A-19	Sequence 19, Appli
37	93	3.2	20	4	US-08-973-961-5	Sequence 5, Appli
38	93	3.2	1381	4	US-09-540-245A-16	Sequence 16, Appli
39	92.5	3.2	669	4	US-09-071-035-264	Sequence 264, App
40	92.5	3.2	1638	4	US-09-071-035-258	Sequence 258, App
41	92.5	3.2	1638	4	US-09-071-035-262	Sequence 262, App
42	92.5	3.2	1638	4	US-09-071-035-266	Sequence 266, App
43	92.5	3.2	2308	1	US-08-015-973-1	Sequence 1, Appli
44	92.5	3.2	2308	1	US-08-448-164-1	Sequence 1, Appli
45	92.5	3.2	2308	4	US-08-081-929-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590, 020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7

Query Match	9.8%	Score 285;	DB 4;	Length 626;
Best Local Similarity	23.8%	Pred. No. 4.7e-20;		
Matches 144;	Conservative 87;	Mismatches 194;	Indels 180;	Gaps 32;
QY	13	DGTSAGQLPEANT-AEPTSMPEVACAATAAATAGOVNIDPWIMNMYQAQGEFTI- 70		
Db	82	DGSITA-----PEQGTWGGVIAEPSAOMSTAADMATKSVDSW-----EAFESFH 128		
QY	71	-----SPNNTPGDILFDLQGLHLPFLSLHAQMYNGWGNMKYKVLLAGNAFTAGRIII 125		
Db	129	TSVNNKSTSEQTKILFKQSLGLPLNPLYSLHLAKLYVAMSGSIEVRFSGVFGGKLA 188		
QY	126	SCIPPGF-AQNTISIAQTAFPHVIAVRVLEPIEVDVRLVFNNDNAPTRMLVCM 184		
Db	189	IVVPGGVDVQSTSMLO---YPHVLFDAQVPEVIFCLPDLRSLTYHLSMDTDTTSIVM 245		
QY	185	LY-----TPLRASGSSCTDFPVIAGRYLTCPSDFSEFLVPPNVEQKTRPFSVPLNPLN 240		
Db	246	VYNDLNPYANDYNSSC---IVT--VETKPGDPKFKHLKPPG-----S 285		
QY	241	TLNSRVPS-LI-----KSMYMSRDH-GOMVQFNGRVRTLDGLOQTTPPTSASOLKIRSV 295		
Db	286	MLTHGSPVSDILPKKSSLSLGNRHSWIDTF-----IIRPFV 322		
QY	296	PHANGNGYN-----LTELDSGYHA-----FESPA-PIGPPDL---G 329		
Db	323	FQANRHFDFNOETAGWSTPRFRPISVTITEQNGAKLGIGVATDVIIVPGIDGMPDFTTIPG 382		

QY 330 EC-----DWHM-----EASPTTQNTGDIK-QINVK-----QESAPA 361  
Db 333 ELIPAGDAITNGTNDITTTATGDTADIIKNNTFRGMWICGSLQRAWGDKISNTAF- 441  
QY 362 PHLGTIOAGLSDSVN-TNMI-----AKLGWVSPVSD-----GHRGVDVDPWI 404  
Db 442 -ITATLDGNNKINFCNTIDQSKIVVFQDAHVHKKQAQTSDDTLALLGYTGIGEQAI- 498  
QY 405 PRYGTSLTEAAOLA--PPIYPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPIQEFVTHFN 462  
Db 499 ---GSDRRVVRISTLPETGARG-GNHPIFYKNSIKLGYVIRSDV-----FNS 543  
QY 463 EOAPTRGEAALLHYLDP-----DTHRNLFGEFKLYPEGEFMTCPVNSSGTGPTQLPI 512  
Db 544 QILHTRSRLSLNHLPLPDPSFAVYRIIDNSNGSDFDIDGIDGDF--SFVGVSGFGKLEFPL 601  
QY 513 NGVFV 517  
Db 602 SASYM 606  
RESULT 2  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-4  
Query Match 9.4%; Score 272; DB 4; Length 623;  
Best Local Similarity .22.2%; Pred. No. 9.7e-19;  
Matches 133; Conservative 84; Mismatches 209; Indels 174; Gaps 27;  
QY 7 DAPT-NMDGTSGAQLVPEANTAEPISMEPVAGAATAAATAGQVNMIDPWIMNNYVQAPQ 65  
Db 72 DLPLERLEGGDGSITTEQGTWVGVIAPESQAQMSAAADMATKSVDSW-----E 122  
QY 66 GEFTI-----SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFT 119  
Db 123 AFFSFTSVNMSTSETQGTWVGVIAPESQAQMSAAADMATKSVDSW-----E 122  
QY 120 AGKIIISCIIPPGF-AAONISIAQATMPPHVIADRVLEPIEVPLEDVRNVLHNNDAFT 178  
Db 183 GGLAAIVVPPGIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRSNLYHLMSDSDT 239  
QY 179 MRLVCMY----TPLRAGSSSGTDFEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFV 234  
Db 240 TSLVIMVYNDLIPYANDTNSGC---IVT--VETKPGDPKFKHLLKPPG----- 284  
QY 235 PNLPLNTLSNRVPS-LI-----KSMYRSRHHQMVQ-----FQ-NGRVTLDDGQLQ-- 279  
Db 285 -----SMLTHGVSVPDLIPKSSLSLGNRHSWSDITFIIRPFVQANRHFDFNQETAGWS 339  
QY 280 -----TTFTSASQCKI-----RGSVFHANGNGN 305  
Db 340 TPRFRPITITVSESNKSLGIGVATDVIVPGIPDGWPDITTIPEQLTPAGIYSITASNGTD 399

QY 306 LTLEDGSPYHAFSPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345  
Db 400 ITTAAG--YDAAETIVNTTNEKSMYICGSLQRAWGDKISNTAFITTAVRKGNISIEPSNT 457  
QY 346 GDVIKQINVKQESAFAPHLGTTQADGLSDSVSVNTNMIAKLGWVSPVSDGHRGVDVP- 401  
Db 458 IDMTKLIV-VYODA-----HVG-----EEVQTSDIITLALLGYTGIGEEAIGSDRQKVRI 505  
QY 402 WILPRYGSITLTEAAQLAPPIYPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPIQEFVTHFN 461  
Db 506 SVLPETGAR-----GNHFIYKNSIKLGYVIRSDV-----FN 539  
QY 462 NEOAPTRGEAALLHYLDP-----DTHRNLFGEFKLYPEGEFMTCPVNSSGTGPTQLPI 511  
Db 540 QILHTRSRLSLNHLPLPDPSFAVYRIIDNSNGSDFDIDGIDGDF-----SFVGVSNLP 592  
RESULT 3  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-2  
Query Match 9.3%; Score 269; DB 4; Length 623;  
Best Local Similarity 29.8%; Pred. No. 2e-18;  
Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;  
QY 7 DAPT-NMDGTSGAQLVPEANTAEPISMEPVAGAATAAATAGQVNMIDPWIMNNYVQAPQ 65  
Db 72 DLPLERLEGGDGSITTEQGTWVGVIAPESQAQMSAAADMATKSVDSW-----E 122  
QY 66 GEFTI-----SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFT 119  
Db 123 AFFSFTSVNMSTSETQGTWVGVIAPESQAQMSAAADMATKSVDSW-----E 122  
QY 120 AGKIIISCIIPPGF-AAONISIAQATMPPHVIADRVLEPIEVPLEDVRNVLHNNDAFT 178  
Db 183 GGLAAIVVPPGIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRSNLYHLMSDSDT 239  
QY 179 MRLVCMY----TPLRAGSSSGTDFEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFV 234  
Db 240 TSLVIMVYNDLIPYANDTNSGC---IVT--VETKPGDPKFKHLLKPPG----- 284  
QY 235 PNLPLNTLSNRVPS 249  
Db 285 -----SMLTHGVSVP 294  
RESULT 4  
US-09-590-020-6  
; Sequence 6, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-6

Query Match 9.1%; Score 262.5; DB 4; Length 622;  
Best Local Similarity 22.3%; Pred. No. 8.9e-18;  
Matches 134; Conservative 82; Mismatches 213; Indels 171; Gaps 28;  
QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPHIMNYYVQAPQ 65  
Db 72 DLPLRLEGDDGSIITPEQGTWGGVIAEPAQMSAADMATGKSYDSEW-----E 122  
QY 66 GEFTI-----SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFT 119  
Db 123 AFFSFHTSVNMSTSETQGKILFKQSLGLPLNPLYLHLAKLYVALAGSVYRFSISGSGVF 182  
QY 120 AGKIIISCIPPGF-AAONISIAQTMFPHVIADRVRLVLEIEVPLEDVRNVLFNNDNAPT 178  
Db 183 GGLAAIIVWPPGIEPVQSTSMQ---YPHFLDARQVEPVIFAIPDLRSNLYHLMSDSTD 239  
QY 179 MRLVCMLY----TPLRASGSSGTDPEVIAGRVLTCPSPDFSELFVPPNVEQKTFPSV 234  
Db 240 TSLVIMVYNDLNPYANDTNSSGC---IVT--VETKPGDPFKFHLKPPG----- 284  
QY 235 PNLPLNTLSNRVPS-LI---KMWVSRDRHGMVQ-----FQ-NGRVTLDGQLQ-- 279  
Db 235 ----SMLTHGSPDLDLPKSSSLWGNRHWSDDITDFIIRPFVQANRHDFDQETAGWS 339  
QY 230 -----TTTISASOLCKI-----RGSVFHANGNGYN 305  
Db 340 TPRFRPTITVSESNMKSGLIGVATDVIYVPGDPGDDTTIPEQLTPAGIYSITASNGTD 399  
QY 306 LTLDGSGPYHAFESPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345  
Db 400 ITTAAG--YDAETIVNTFKSMYICGSLQRAWDKKSINTAFITTAVRKGSNIEPSNT 457  
QY 346 GDVTKQINVKQESAFAPHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWV-I 404  
Db 458 IDMTKLV-VYQDA---HVG-----EEVQTSIDTLALLGYTGIGEEAIGSDRDKVRI 505  
QY 405 PRYGSTLTEAAQLAPPYIPGFGAEIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQ 464  
Db 506 SVLGETGARG-----GNHPIFYKNSIKLGVIRSIDV-----FNSQI 542  
QY 465 APTRGEAALHY-LDPDTHR-----NLGEFKLYPEGFMTCVPNSSGTGQTLTP 511  
Db 543 LHSSRQLSLNLYLLPPDSFAVYRILDSNGSWFDIGIDTEFPDGF-----SFGVSNLP 595

RESULT 5  
US-07-852-260-2  
; Sequence 2, Application US/07852260  
; Patent No. 5525715  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/852,260  
; FILING DATE: 19920619  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-852-260-2

Query Match 5.1%; Score 149; DB 1; Length 2206;  
Best Local Similarity 20.3%; Pred. No. 2.5e-05;  
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSCAGOLVPEANTAEPISEMPVAGATAAATAGQV-NM-----IDPWNNYYVQAPQ-- 66  
Db 347 TFGSNYLTSDNHQSCAI--PEFDVTPDIPGEVNMMLAEIDTWIPLNLESTKRNTM 405  
QY 67 ---EFTISPNTPGDILFDLQGLPHLNPFLSH-----LAQMYNGWGNKVKVLLAGNAF 118  
Db 406 DMVRYTLSDSADLSQPLCLSLSPAFDPLRSLHTMLGEVLNYYTHWAGSLKFTFLCGSMM 465  
QY 119 TAGKIIISCIPPGFAONISIAQTMFPHVIADRVRLVLEIEVPLEDVRNVLPH--NNDNA 176  
Db 466 ATGKILVAYAPG-AQPPTSRKEAMLGTHVINDLGLQSSCTMVYVPWISNVTYRQTQDSF 524  
QY 177 PMRLVCMLY-----TPLRASGSSGTDPEVIAGRVLTCPSPDFSELF--V 221  
Db 525 TEGGYISMFYQTRIVVYPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTHISQSAL 576  
QY 222 PNVEQKTK-----PFSVPNLP-----LNTLSNRVPSLIKSNM 255  
Db 577 PCGIEDLTSEVAGCALTLSPKQDLSLDPOTKASGPAHSKEVPALTAVETGATNPLAPSDT 636  
QY 256 VSRDHGMVQ-----FQNGRVTLDGQLQGTTFPTSASOLCKIRGSVPH----- 297  
Db 637 VCTRH--VQRRSRSESTIESFARGACVAILVEVDNEQPTTFAQKLFAMWIRITYKDTVOL 694  
QY 298 -----ANGNGYNLTEDGSPYHAFESPAIGPDPDLGECOW 333  
Db 695 RRKLEFFTYSRFDEMTFVVVTANFTNANNGHALNQVYQIMYIPPGAPTPKSHDDY---TW 751  
QY 334 HMEASPTTQFNTGDIKQINVKQESAFAPHLGTIOA-----DGLSDVSVNTNMIKLG 386  
Db 752 QTSNPSIFITYGAAPARISV-----PYVGLANAYSHFYDGFAGKVPKLTQANDQIG 802

RESULT 6  
US-08-461-503-2  
; Sequence 2, Application US/08461503  
; Patent No. 5834302  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent

```

; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-503-2

```

```

Query Match          5.1%; Score 149; DB 2; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGQLVPEANTAEPISEMPVAGATAATAAGOV-NM-----IDPWIMNNYVQAPOG-- 66
DB 347 TPGSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ---EFTSPNNTPGDILFDLQGLPHLNFSLH-----LAQMYNGWGNMKVKVLLAGNAF 118
DB 406 DMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIISCIPGFAAQNISTAQATMFPHVIAADVRLVLEPIEVLNVLPH--NNDNA 176
DB 466 ATGKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPWISNVTYRQTQDSF 524
QY 177 PTMLRVLCMLY-----TPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFL-----V 221
DB 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTHISQSAL 576
QY 222 PPNVEQKTK-----PFSVNPILP-----LNTLSNRVPSLIKSM 255
DB 577 PQGIEDLTSEVAGALTLSPKQDLSLPTDKASGPAHSKEVPALTAVETGATNPLAPSDT 636
QY 256 VSRDHQOMVQ-----FQNGRVTLDGQLGTTPTTSASOLCKIRGSVFH----- 297
DB 637 VQTRH--VVQRRSRSESTIESFFARGACVAILIEVDNEQPTTTRAKLFAWMRIYKDTVOL 694
QY 752 QTSSNPISFYTYGAAPARISV-----PYVGLANAYSHFYDGFPAKVPKLTDANDQIG 802

```

```

RESULT 7
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-250-2

Query Match          5.1%; Score 149; DB 4; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGQLVPEANTAEPISEMPVAGATAATAAGOV-NM-----IDPWIMNNYVQAPOG-- 66
DB 347 TPGSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ---EFTSPNNTPGDILFDLQGLPHLNFSLH-----LAQMYNGWGNMKVKVLLAGNAF 118
DB 406 DMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIISCIPGFAAQNISTAQATMFPHVIAADVRLVLEPIEVLNVLPH--NNDNA 176
DB 466 ATGKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPWISNVTYRQTQDSF 524
QY 177 PTMLRVLCMLY-----TPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFL-----V 221
DB 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTHISQSAL 576
QY 222 PPNVEQKTK-----PFSVNPILP-----LNTLSNRVPSLIKSM 255
DB 577 PQGIEDLTSEVAGALTLSPKQDLSLPTDKASGPAHSKEVPALTAVETGATNPLAPSDT 636
QY 256 VSRDHQOMVQ-----FQNGRVTLDGQLGTTPTTSASOLCKIRGSVFH----- 297
DB 637 VQTRH--VVQRRSRSESTIESFFARGACVAILIEVDNEQPTTTRAKLFAWMRIYKDTVOL 694
QY 298 -----ANGNGYNLTLDGSPYHAFESPAPIGFPLDGCWD 333

```

```
Db 695 RRLKFFTSYREDMEFTVVTANETNANNGHALNQVQIMVPPGCAPTPKSWDDY---TW 751
QY 334 HMEASPTTQFNTGDIKQINVKQESAFAPHLGTTQA-----DGLSDVSVNTNMIATKL 386
Db 752 QTSNPSIFTYGAAPARISV-----PYVGLANAYSHFYDGFAPKPLKTDANDQIG 802

RESULT 8
US-09-423-890-2
; Sequence 2, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPPC
; CURRENT FILING DATE: 2000-03-06
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-890-2

Query Match 3.7%; Score 106; DB 4; Length 1302;
Best Local Similarity 18.7%; Pred. No. 0.24;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

QY 1 MMASKDAPTNMDGTSGAGQLVPENTAEPISEMPVAG--ATAAATAGQVNMIDPWIMNN 59
Db 400 LLLANGESTNGSGGSG--GSL--SAGAASGSGSQPSISGDVVEACCSVLISVCADP-VYKV 455
QY 60 YVOA---PQGEFTISPNNTPGD-ILFDLQLGPHLNPF------SHLAQMY 100
Db 456 YVAALKTLRLMVLVYTPCHSLAERIKLQRLRPVVDITLVKCADANSRTSLSISVLEIC 515
QY 101 NGWGNMKV--KVLLAGNAFTAG-KIIISCIPPGFAAQN-----ISIAQA 142
Db 516 NQAGKLVAGREILKAGSIGVGGDYVLSCILGNAESNNQWELLGRCLLIDRLLEPPA 575
QY 143 TMFPHVIA-DVRVLEPIEVLDPVNRVL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196
Db 576 EFYPHIVSTDVSAEPVEIRYKLLSLTLFALQSIDNSHSMVGLSRRIY-----625
QY 197 GTDPPVIAGRVLTCPSPDFSLFVPPNVEQTKPFVNPPLNTLSNRVPSLIKSMV 256
Db 626 -----LSSARWTVAPVAFSKLVTM-----LNASGSTHTRMRRRLMA 663
QY 257 SRDHQMVQ-FQNG-RVTLDG---QLQGTTPTSASQLCKIRGSVFHANGNGYNLTELD- 310
Db 664 IADEVEIAEVIQLGVEDTVGDHQSLOALPASCLSENSLEHTVHREKTGKLSATRLSA 723
QY 311 -----GSPYHAFESPAPIGFPDGLGECDDHMEAS 338
Db 734 SSEDISDLRAGSVGLPSSTTTEQPKPAVQTKGRPHSQCLNSSPLSHAQL-----MPPA 777
QY 339 PTFQNTGDVIKQINVKQESAFAP 362
Db 778 PSAPCSSAPSVDPDISKRPQAFVP 801

RESULT 9
US-09-423-890-8
; Sequence 8, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extracellular Matrix
; FILE REFERENCE: CPI-004DVCP3
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: US/08/628,829A
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
```

```
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPPC
; CURRENT FILING DATE: 2000-03-06
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1493
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-423-890-8

Query Match 3.7%; Score 106; DB 4; Length 1493;
Best Local Similarity 18.7%; Pred. No. 0.31;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

QY 1 MMASKDAPTNMDGTSGAGQLVPENTAEPISEMPVAG--ATAAATAGQVNMIDPWIMNN 59
Db 591 LLLANGESTNGSGGSG--GSL--SAGAASGSGSQPSISGDVVEACCSVLISVCADP-VYKV 646
QY 60 YVOA---PQGEFTISPNNTPGD-ILFDLQLGPHLNPF------SHLAQMY 100
Db 647 YVAALKTLRLMVLVYTPCHSLAERIKLQRLRPVVDITLVKCADANSRTSLSISVLEIC 706
QY 101 NGWGNMKV--KVLLAGNAFTAG-KIIISCIPPGFAAQN-----ISIAQA 142
Db 707 KQAGELAVGREILKAGSIGVGGDYVLSCILGNAESNNQWELLGRCLLIDRLLEPPA 766
QY 143 TMFPHVIA-DVRVLEPIEVLDPVNRVL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196
Db 767 EFYPHIVSTDVSAEPVEIRYKLLSLTLFALQSIDNSHSMVGLSRRIY-----816
QY 197 GTDPPVIAGRVLTCPSPDFSLFVPPNVEQTKPFVNPPLNTLSNRVPSLIKSMV 256
Db 817 -----LSSARWTVAPVAFSKLVTM-----LNASGSTHTRMRRRLMA 854
QY 257 SRDHQMVQ-FQNG-RVTLDG---QLQGTTPTSASQLCKIRGSVFHANGNGYNLTELD- 310
Db 855 IADEVEIAEVIQLGVEDTVGDHQSLOAVAPTSCLSENSLEHTVHREKTGKLSATRLSA 914
QY 311 -----GSPYHAFESPAPIGFPDGLGECDDHMEAS 338
Db 915 SSEDISDLRAGSVGLPSSTTTEQPKPAVQTKGRPHSQCLNSSPLSHAQL-----MPPA 968
QY 339 PTFQNTGDVIKQINVKQESAFAP 362
Db 969 PSAPCSSAPSVDPDISKRPQAFVP 992

RESULT 10
US-08-628-829-4
; Sequence 4, Application US/08628829A
; Patent No. 633170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extracellular Matrix
; FILE REFERENCE: CPI-004DVCP3
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: US/08/628,829A
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
```

```
; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-628-829-4

Query Match          3.7%; Score 106; DB 4; Length 1593;
Best Local Similarity 18.7%; Pred. No. 0.34;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

QY 1 MMASKDPTNMDGTSGAGOLVPEANTAEPISEPVAG-AATAAATAGOVNMIDPIMNN 59
Db 691 LLLANGESTNGSGGSGSL--SAGAAAGSSQPSISGDVVEACCSVLSTVCADP-VYKV 746

QY 60 YVQA---PQGEFTISPNTPGD-ILFDLQGLPHLPFL-----SHLAQMY 100
Db 747 YVAALKTRAMLVYTPCHSLAERIKLQRLRPVVDITLVKCADANSRTSOLSISTVLELC 806

QY 101 NGWGNMKV--KVLLAGNAFTAG-KIISCIPPGFAQN-----ISIAQA 142
Db 807 KQAGEIAVAREILKAGSIGVGVDYVLSCLIGNQAESNNWQELLRLCLIDLRLLEFPA 866

QY 143 TMFPHVIA-DVRVLEPTEVLEPVRNVL-----FHNDNAPTMM--RLVCMLYTPLRASGSS 196
Db 867 EFYPHIVSTDVQAEPVEIRYKLLSLTALFALOSIDNSHSMVGKLSRIY----- 916

QY 197 GTDFPVTAGVLTCPDPSFELFVPPNVEQTKPFSVNPMLPLNTLSNRKVPSLIKSMV 256
Db 917 -----LSSARMVTAAPVAFSKLVTM-----LNASGSTHFTMRRLMA 954

QY 257 SRDHGQWQ--FQNG-RVTLDG---QLQGTPTTSASOLCKIRGVSFVHANGNGYNLTELD- 310
Db 955 IADEVEIAEVIQLGVETVDGHDQSLQAVAPTSCLENSLEHTVHREKTKGLSATRLSA 1014

QY 311 -----GSPYHAFESPAPIGFPDGLGCDNHRMEAS 338
Db 1015 SSEDISRLAGSVGLPSSTTTEQPKPAVQTKGRPHSQCLNSSLPLSHAQL-----MFP 1068

QY 339 PTTQFNTGDIKQINVKQESAFAP 362
Db 1069 PSAPCSSAPSPDLSKRPQAFVP 1092

RESULT 11
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-09-091-219-24

Query Match          3.7%; Score 106; DB 4; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.64;

Matches 63; Conservative 31; Mismatches 113; Indels 76; Gaps 13;

QY 5 SKDAPNTNMDGTSGAGOLVPEANTAEPISEME-----PVAGAATAAATAGOVNMIDPIMNNY 60
Db 270 SKLASSAFSGLEGA--LLADKKTEETITLEDRLITRNGHTTSTTQSSGV-----TY 320

QY 61 VOAPQGEFTISPNTTPGDI-----LFDLQLGP-----HL-----NPF 92
Db 321 GYATAEDFVSGDNTSGLETRVVQAEERFFKTHLFDWVTSDFGRCHLLELPTDGHKGVYCSL 380

QY 93 LSHLAQMYNGWGNMKVKKVLLAGNAFTAGKIIISCIPGFAAQNISIAQATMFPHVIADV 152
Db 381 TDSYAYMRNGW---DVEVTAVGNQFNGGCLLVAMVPPELYSTQKRELYQLTLFPHQFINP 436

QY 153 R--VLEPTEVPLEDYRNV-----LFHNDNAPTMRVLCMLY--TPLRA 191
Db 437 RTNMTAHTVTPGVGNRYDQYKVKHPWTLLVMVAVPLTVNTGAPQIKVYANIAPTNVHV 496

QY 192 SG---SSSGTDPFVIA---GRVLTCP---SPDFSFLFLVPPN 224
Db 497 ACEFTPSKEGIRPVACSDGVGGLVTTDPKTADPVYGVKVFNP 539

RESULT 12
US-09-749-588-2
; Sequence 2, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEOTIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-749-588-2

Query Match          3.6%; Score 103.5; DB 4; Length 1170;
Best Local Similarity 20.9%; Pred. No. 0.37;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

QY 9 PTNMDGTSGAGOLVPEANTAEPISEMPVAGAT-AAATAGOVNMIDPIMNNYVOAPOGE 67
Db 633 PVRMD---NAVPIVPQAPAAQLQIQ--SGVLTOGSCPTLMVATLHPQVA----- 677

QY 68 FTISPNTTPGDIILFDLQGLPHLPFLSLHAQMYNGWGNMKVKKVLLAGNAFTAGKIIISC 127
Db 678 -TITPQYA---VPFTLSCAAGRPALVEQTAVLQNPQGTQ-QILLPS---TWQOL----- 725

QY 128 IPPGFAQNISIAQATMFPHVIADVRLVLEPTEVPLEDYRNVLFHNDNAPTMRVLCMLYT 187
Db 726 --PGVALHN-SVQPTAMIPAMGSGQ-----QLADWRNAHSHGNOYSTIMQPSLLTN 775

QY 188 -----PLRAS-----GSSSGTDPFVIAGRY--LTCPS 213
Db 776 HVTLATAQPLANVGVAHVVRQOQSSSLPSKKNQKQSAAPVSSKSLD--VLPSQVYSLVGSSP 833

QY 214 ---DFSFLFLVPPNVEOKTKPSPVNLPLNTLS-----NSRVPSLIKSMVMS 257
Db 834 LRTTSSYNLVP--VQDQHQIIPIDTPSPVSVTITRSDTDEEDNKKYKES--SSGLKP 889

QY 258 RDHGQWQFQNGRVTLDGLOGTTPTSASOLCKIRGVSFVHANGNG---GYNLTELDSGPY 314
Db 890 RSN--VISYVTVNDSPDSDSLSSPYSTDTLSALRG-----NSGSLVLEGPRVAVDGTGT 942

QY 315 HAFESPAIGFPDGLGCDNHRMEASPTTQFNTGDIKQINVKQESAFAPHLGTIQADGLS- 373
```

Db 943 RTIIVP-PLK-TQLGDCVTATQASGLLSNKTKPVASVSQSSGCCITPTGTYRAQRGGTSA 1000  
QY 374 -----DVSNTNMTAKLGWSPVSDGHRGDVDPWIPRYGTLTEAAQLAPPYPPGFGGEA 429  
Db 100- AOPNLNLSONQSSA-----APTSQERSNPAP-----RQQAFAVAP-----LSQA 1040  
QY 430 IVFMSDFPIAHGTNG-----LSVPCTIPQEFVTHFVNEQAPTGCEA-----ALLHYLDPD- 480  
Db 1041 PYTFQHGSPD-HSTGPHPLAPAPHLPSQ--AHLYTYAAPTSAALGTSSTAHLESPOG 1097  
QY 481 THRNLFGEFKLYPEGFMTCPVNSGSGTGPOTL 510  
Db 1098 SSRHAAAYTHPSTLTHQVPVS--VGPSLL 1125  
RESULT 13  
US-08-460-860-4  
; Sequence 4, Application US/08460860  
; Patent No. 5665584  
; GENERAL INFORMATION:  
; APPLICANT: HATAMOTO, OSAMU  
; APPLICANT: WATARAI, TERUO  
; APPLICANT: MIZUSAWA, KIYOSHI  
; APPLICANT: NAKANO, EIICHI  
; TITLE OF INVENTION: A DNA FRAGMENT CONTAINING A TANNASE  
; TITLE OF INVENTION: GENE, A RECOMBINANT PLASMID, A PROCESS FOR PRODUCING  
; TITLE OF INVENTION: TANNASE, AND A PROMOTER  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,860  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 159973/1994  
; FILING DATE: 12-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 83973/1995  
; FILING DATE: 10-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F  
; REFERENCE/DOCKET NUMBER: 7127-001-0X  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 588 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-460-860-4  
Query Match 3.5%; Score 102.5; DB 1; Length 588;  
Best Local Similarity 21.7%; Pred. No. 0.15;  
Matches 66; Conservative 41; Mismatches 100; Indels 97; Gaps 18;  
QY 273 LDGQLQGGTTPTSASOLCKIRGSVHANGNGYNLTDLGDSYH-AFESPAPIGFDPDGEC 331  
Db 264 LDGRDGV--VSRDLCKL-----NENLTSTIGEPYCAAGTSTSLGF----- 304

QY 332 DHMEASPTTQNTGDVLIKQINVKOE-----SAPAP-HLGTIOADGLSDVSV-----N 378  
Db 305 -----GFSNG-----KRSNKRQAEGSTSYQPAQNGTGTARGVAVAAQAIYDGLHNS 351  
QY 379 TMMIAKLGW--VSPVSDG---HRGDVDPWV---IPRYGS-----TLTEAAQLAPPY 421  
Db 352 KGERAYLSWQIASELSDAETENSDGTGWELNIPSTGGGYVTKFIQLNLNLDNLNNVT 411  
QY 422 YPPGGEAIVFFMSDFPIAHGTNGLSVPTIQQ-----EFVTHFVNEQAPTGCEAAL 473  
Db 412 Y-----DTLVDWMTGMVRYMD---SLQTTLPDLTPFOSSGGKLLHYHGESDPSIPAASS 463  
QY 474 LHYLDP-----DTHRNLFGEFKLYPEGFMTCPVNSGSGTGPOTLPIGVFVF 518  
Db 464 VHYWQAVRSVMYGDKTTEEALEALEWDYQFYLLP-GAAHCGTNSLQPGP--YPENNM 520  
QY 519 VSKV 522  
Db 521 IDWV 524  
RESULT 14  
US-09-171-699-10  
; Sequence 10, Application US/09171699  
; Patent No. 6448389  
; GENERAL INFORMATION:  
; APPLICANT: The Wistar Institute of, Anatomy & Biology  
; Gonczol, Eva  
; Berencsi, Klara  
; Kari, Csaba  
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and  
; Uses Therefor  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Center, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171,699  
; FILING DATE: 19-Jan-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/015,717  
; FILING DATE: 23-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kodroff, Cathy A.  
; REGISTRATION NUMBER: 33,980  
; REFERENCE/DOCKET NUMBER: WST66APCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1048 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-171-699-10  
Query Match 3.5%; Score 102; DB 4; Length 1048;  
Best Local Similarity 21.8%; Pred. No. 0.44;  
Matches 120; Conservative 50; Mismatches 214; Indels 166; Gaps 25;

QY 15 TSGAGQLVPEANTARPISMEPV-----AGAATAAATAGQVNMIDPWNMNNVVOAPQG 66  
 Db 578 TFGAGVNVNVPAGAAAILTPTVNESTAPAPAPTPTFACTQTPVNGNSPW-----APTA 630  
 QY 67 EF--TISPNTTGGDLFDLQGLPHL--NPFSLHQAQYNGWGNMVKVLLAGNAFTAGK 122  
 Db 631 PLPGDMNANPREAWALK--NPHLAYNPFMPPTS-----TASQ 669  
 QY 123 IISCIPGFAAQNTSIAQATMFHVIADRVL--EFIEVPLEDVNRVLFHNNDAFTMR 180  
 Db 670 NVTSTPRPSTPRAAVTQ--TASRDAADDEVWALRDQTAESPVEDSE-----714  
 QY 181 LVCMLYTLRASGSSGSDPFIAGRVLTCPSPDFSLFLVPPNVEOKTKPESVP-----235  
 Db 715 -----EEDDDSSDTGVSUVSLGH--TTPSSDYNNDVISPPS---QTPEQSTPSRIK 760  
 QY 236 ----NPLNTLSNSRVPSLIKSMVSRDH-----GOMVQFQNGRVTLDCQLQGTPTTSASQL 288  
 Db 761 AKLSSPMTTTSTQSKPVLGK--RVATPHASARAQTVTSTPVQGRLEKQVSGTPTVPATL 818  
 QY 289 CKIRGSVHANGNGNYLTLDGSPYHAFESPAPIGPDLGCECDHMEASPTTQFNT---345  
 Db 819 LQ-----PQP-----ASSKTTSSRNVT 836  
 QY 346 -GDVVKOINVKESAFAPHLGTIQADGLSDYSVNTNM-----AKLGVWSPVSD--GHR 396  
 Db 837 GAGTSSASARQPSASASVLSPTEDDVVSPATSPLSMLSSASPSAKSAPPSPVKGRGRS 896  
 QY 397 GDVDPWVPRVGTSTEARAQLAPPIYP-----PFGGAIVFFMSDFPIAGHGLSVPCPT 451  
 Db 897 VGV--PSLKPTLGG-----KAVGRRPSPVPSVSGAPRLSGSSRAASTTPTPAVTTVPYPPSS 952  
 QY 452 IQPEFVTHFVNQAAPT---RGEAALLHYLDPTDTHRNLECFK-----LYPEGFMT 498  
 Db 953 TAKSSVSNAPPVSPSILKPGASAL-----QSRSTGTRAAGSPVKSTTGKMTVAFDLS 1007  
 QY 499 VPNSSGTGPQ 508  
 Db 1008 SPQSGTGPQ 1017

RESULT 15  
 US-09-615-192A-349  
 ; Sequence 349, Application US/09615192A  
 ; Patent No. 6410718  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bloksberg, Leonard N.  
 ; APPLICANT: Havukala, Ilkka  
 ; TITLE OF INVENTION: Materials and Methods for the  
 ; TITLE OF INVENTION: Modification of Plant Lignin Content  
 ; FILE REFERENCE: 11000.100364U  
 ; CURRENT APPLICATION NUMBER: US/09/615.192A  
 ; CURRENT FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 08/975,316  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: US 08/713,000  
 ; PRIOR FILING DATE: 1996-09-11  
 ; PRIOR APPLICATION NUMBER: US 09/169,789  
 ; PRIOR FILING DATE: 1998-10-09  
 ; NUMBER OF SEQ ID NOS: 405  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 349  
 ; LENGTH: 544  
 ; TYPE: PRT  
 ; ORGANISM: Pinus radiata  
 US-09-615-192A-349

Query Match 3.5%; Score 100; DB 4; Length 544;  
 Best Local Similarity 19.9%; Pred. No. 0.24;  
 Matches 109; Conservative 68; Mismatches 197; Indels 174; Gaps 30;  
 QY 7 DAPTNMDGTSGAGQLVPEANTARPISMEPVAGAAATAAATAGQVNMIDPWNMNNVVOAPQG 66

Db 42 DRPCVINGATORTTYAE---VELISRRVSAGNLGVGGQDVIML---LLQN---CPEF 92  
 QY 67 EF-----TISPNTTGGDI-----LFDLQGLPHNPFISHLAQMYN 101  
 Db 93 VFAFLGASYRGAISTANPEFTVTCGEIAKQASAAKAKIVITQAAAFADKVRFAEE---N 147  
 QY 102 GWVGNMVKVLLAG-----NAFTAGKI---TISCIPPGFAAQNTSIAQ 141  
 Db 148 G-----VKVVCIDTAPEGCLHFSELMOADENAAPADVPDDVLALP--YSSGTTGLPK 199  
 QY 142 ATMFPHVIADRVLEPIEVPLEDVNRVLFHNNND---NAPTMRLVCM---LYTPLRASGS 194  
 Db 200 GYMLTHRGQVTSVAQVD---GDNPNLYFFHKEDVILCTPLPLFHIYSLSNVMFCAALRVGAA 256  
 QY 195 SSGTDFFIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPNPLPLNTLSNSRVPSLIKSM 254  
 Db 257 ILMOKFEIVALMELVQRYRYVTILPIVPPIVLEIAKSAEVDYDL-----SSIRTI 307  
 QY 255 MYSRDHGQMVQFQNGRVTLDCQLQGTPTTSASQLCKIRGSVHANGNGNYLT---LD 310  
 Db 308 M-----SGAAPMKRELEDT-----VRAKLPNAKLGQGYGMTAGPVLA 345  
 QY 311 GSPYHAFESPAPIGPDLGCECDHMEASPTTQFNTGDIKQINVK-----QESAPAP--HL 364  
 Db 346 MCPAPAKE--PEEI---KSGAC-----GTVVRNAEMKIYDPTETGASLPRNQA 387  
 QY 365 GTIQADG-----LSDYSVNTNMIAKLGWVSPVSDGHRGDV-----DPWVTPRYGST 410  
 Db 388 GEICIRGHQIMKGYLNDAAEATANTIDKEGWL-----HTGDIGYIDDDDELFIYDRLKEL 441  
 QY 411 LT-EAAQLAPP-----IYPPFGGAIVFFMSD-----FPIAH--GTNGLSVPTCTPQEF 456  
 Db 442 IKYKGFQVAPAELEAMLIAPHISDAAVVPMKDEVAGEVPVAFVVKSN-----SVITEDE 497  
 QY 457 VTHFVNEQ 464  
 Db 498 IKQYISKQ 505

Search completed: March 10, 2003, 19:00:44  
 Job time : 26 secs



GenCore version 5.1.3  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:25:00 ; Search time 23 Seconds  
(without alignments)  
2277.968 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1974.5	68.2	546	2 B37491	major capsid prote
2	1946	67.2	530	2 B37471	capsid protein - N
3	1223	42.2	542	2 S60616	capsid protein - h
4	1151.5	39.8	539	2 S40111	capsid protein - h
5	316	10.9	2344	1 RWRWH	genome polyprotein
6	315	10.9	576	2 A53982	capsid protein - E
7	311.5	10.8	2344	2 S55399	genome polyprotein
8	307.5	10.6	2344	2 S64740	genome polyprotein
9	300.5	10.4	702	1 A48562	coat protein - San
10	277.5	9.6	668	2 JQ2354	capsid protein - f
11	276	9.5	671	1 VCMWF9	coat protein - fel
12	274	9.5	668	1 VCMWFF	coat protein - fel
13	269	9.3	668	1 VCMWFF	coat protein - fel
14	264	9.1	703	1 C48562	coat protein - San
15	261.5	9.0	668	2 JQ2356	capsid protein - f
16	147	5.1	2206	2 S03822	genome polyprotein
17	146.5	5.1	2205	1 GNNY2W	genome polyprotein
18	145.5	5.0	2207	1 GNNY5P	genome polyprotein
19	145	5.0	2206	1 GNNY4P	genome polyprotein
20	143.5	5.0	2194	2 S09553	genome polyprotein
21	141.5	4.9	2207	1 GNNY27	genome polyprotein
22	139.5	4.8	2206	1 GNNY27	genome polyprotein
23	137.5	4.7	2209	1 GNNY2P	genome polyprotein
24	136.5	4.7	2209	1 GNNY2P	genome polyprotein
25	135.5	4.7	2207	1 GNNY1P	genome polyprotein
26	133.5	4.6	2179	1 GNNY4A	genome polyprotein
27	128.5	4.4	613	2 T35828	acetolactate synth
28	125	4.3	3473	1 A46112	genome polyprotein
29	125	4.3	3473	2 S27927	polyprotein - rice

30 124 4.3 2214 1 A48548 genome polyprotein  
31 123.5 4.3 2175 1 GNNYBE genome polyprotein  
32 122.5 4.2 3085 2 T00327 polyprotein - infe  
33 122 4.2 3263 2 E82410 hypothetical prote  
34 117.5 4.1 2185 1 JQ2021 genome polyprotein  
35 117 4.0 2183 1 GNNYB4 genome polyprotein  
36 116.5 4.0 2206 1 GNNY21 genome polyprotein  
37 116 4.0 1037 2 T13350 genome polyprotein  
38 115 4.0 2185 1 GNNYSV genome polyprotein  
39 115 4.0 2185 1 GNNYSH genome polyprotein  
40 113.5 3.9 2164 1 GNNY89 genome polyprotein  
41 113 3.9 833 2 S45041 genome polyprotein  
42 111.5 3.9 516 2 S28060 serum response fac  
43 111 3.8 3624 2 AD0835 large repetitive p  
44 110.5 3.8 5188 2 B85547 probable RTX famil  
45 110.5 3.8 5291 2 F90696 hypothetical prote

#### ALIGNMENTS

RESULT 1

B37491

major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C>Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000

C/Accession: B37491

R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Science 259, 516-519, 1993

A>Title: Sequence and genome organization of a human small round-structured (Norwalk-

A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A:Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LAM>

A:Cross-references: GB:L07418; NID:q1236787; PIDN:AAA92984.1; PID:q295114

A>Note: sequence extracted from NCBI backbone (NCBI:P123458)

A>Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, seroty

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 68.2%; Score 1974.5; DB 2; Length 546;

Matches 357; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPWIMNNY 60

Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGTTAVATAGVNMIDPWVNF 60

Qy 61 VQAPQGETTISPNTPGDILFDLQGLHNPFLSHLAQMYNGWGNMVKVLLAGNAFTA 120

Db 61 VQSPQGETTISPNTPGDILFDLQGLHNPFLSHLSQMYNGWGNMVRILLAGNAFTA 120

Qy 121 GKILISCLIPCPGFAONISIAQTAMPHTVADRVLEPLEVPLEDVRNLFHNDNAPTR 180

Db 121 GKILVCCPPGFTSSSLTAQTALPHVIAVRTLEPLEDVRNLYHTNDNQPTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNPLN 240

Db 181 LVCMLYTPLRGSGSGNSDSFVAGRVLTAPSSDFSLFLVPPPTIEQTRATFTVNPIDQ 240

Qy 241 TLSNSRVPSLIKSMVSRDHQMYQFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHANG 300

Db 241 TLSNSRFPSLIQGLMLSPDASQVVFQFQNGRCLIDQLLGTTPATSGQLFRVRGKI--NQG 298

Qy 301 GNGYNLTLDGSPYHAPESPAPIGFPDLGECDDHWEASPT--TQNTGDVIRKQINVKOB-S 358

Db 299 ARTUNLTETVDGKPFMAFDPVGFPPDFGKCDWHMIRISKTPNNTGSGDPMRSVSQTVNQ 358

Qy 359 AFAPHLGTLQADGLSDVSVENTNMIATKLGWSPVSDGHGRGDVDPVVIPIRYGSLTLEAAQLA 418

Db 359 AFAPHLGTLQADGLSDVSVENTNMIATKLGWSPVSDGHGRGDVDPVVIPIRYGSLTLEAAQLA 418



C:Species: human calicivirus  
A:Variety: strain Bristol isolate B493  
C:Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C:Accession: S40111  
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A:Reference number: S40111  
A:Accession: S40111  
A:Molecule type: genomic RNA  
A:Residues: 1-539 <GRE>  
A:CROSS-references: EMBL:X76716; NID:g436410; PIDN:CRA54134.1; PID:g436411  
A:Experimental source: human enteric calicivirus strain Bristol isolate B493  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match 39.8%; Score 1151.5; DB 2; Length 539;  
Best Local Similarity 44.9%; Pred. No. 9.9e-78;  
Matches 254; Conservative 78; Mismatches 185; Indels 49; Gaps 16;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAETISMEPVAGAATAAATAGOVNMDIPWIMNNY 60  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 MKMASNDANPS-DGS--AANLVPEVNN-EYMALEPVVGAIAAPVAGQNVDPWIRNRF 56  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
QY 61 VQAPQGETTSPNTPGDIILQLGPHLAPFSLHQAQYNGVGNKVKVLLAGNAFTA 120  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 57 VQAPGGETTSPNAPGELIWSAPGLDPLAPYLSHLSRMVNGYAGGFVQVILAGNAFTA 116  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
QY 121 GKIIISICPPGFAAQNISIAQATMFPHVIAADVRLVEIPLEVDVRLVLFH-NNDNAPT 179  
| | | : | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
Db 117 GKVLFAAVPNPFTGELSPSQVTFMFPHIIIVDQLEPVLPLPDVRRNFTHYNOANDSTL 176  
| | | : | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
QY 180 RLVCMLYTLPLRASGSSGTDPPFVIAGRLVLCPSPDFLEFLVPPNVBQKTKPFSVPNLPL 239  
: | : | | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
Db 177 KLIAMLYTLPLRA--NNAGDDVFTVSCRVLTRPSDPDFELVPTVESRTKPTVPVLTV 234  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
QY 240 NTLNSRVPSLIKSMVSRDHQGVQFQNGRVTLDGOLQGTTPTSASQLCKIRGSVFHAN 299  
: | | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
Db 235 EEMSRSFPPILEKLYTGPSSAFVQVQNGRCTDGVLLGTTQLSAVNICNFRGDVTHIA 294  
| | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
QY 300 GGNCY--NLTELDGSPYHAFES-PAPIGFDL-----GECDHMEASPTTQFNT 345  
| : | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
Db 295 GSHDTYNLASQNNWSNYDTEEPAPLGTDFGVCKIOGLITQTRADGSTRAHKAT-VST 353  
| : | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
QY 346 GDVIKQINVAQESAFAPHLGTIQ--ADGLSDSVSVNTNIAKLGWSPVSDG--HRGDVDP 401  
| | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
Db 354 GSV-----HFTKLGSVQFTTNDNFQAGQN--TKFTPVGVIQGDHDIHQNEPQQ 401  
| | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
QY 402 WVIPRYGSTTEAAQLAPPYIPPGFGRAIVFFMSDFPIAHGTNGLSVPCPTIPOGFVTHFV 461  
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :  
Db 402 WLLPNYSGRTGHNVHLAPAVPTPGGQLLFFRSTMPGCGSGYPNMNLLDCLLPQEWLHFY 461  
| | | | | : | | : | | | | | : | | | | | : | | | | | : | | : | | :  
QY 462 NEQAPTGEALLHYLDPDTHNLGEKLYPEGMTCVPSNSSGTGPQ--TLPINGVFEV 519  
| | | : | | : | | | | | : | | : | | | | | : | | | | | : | | : | | :  
Db 462 QEAAPQSDVALLRFVNPDTGRVLFECKLHKSGYITV----AHTGPYDLVLPNGYPRFD 517  
| | | : | | : | | | | | : | | : | | | | | : | | | | | : | | : | | :  
QY 520 SWVSRYQLKPVGTAGPACRLGIRRS 545  
| | | : | | : | | | | | : | | : | | | | | : | | | | | : | | : | | :  
Db 518 SWVNOFTLAPMGNG-----TGRERRA 538  
| | | : | | : | | | | | : | | : | | | | | : | | | | | : | | : | | :

RESULT 5  
RMWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Meyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557; PMID:1840711  
A:Accession: A41039

A:Molecule type: genomic RNA  
A;Residues: 1-2344 <MEY>  
A;Cross-references: GB:M67473  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: nucleotidytransferase

Query Match 10.9%; Score 316; DB 1; Length 2344;  
Best Local Similarity 28.3%; Pred. No. 1.le-14;  
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

QY 12 MDGTSAGQLVPEANTAEPTIS-----MEPVAGAATAA-----ATAG----- 47  
| : | : | : | : | : | : | : | :  
Db 1766 MESKARAAPGEAAAGTATTASPGCTTTDGMDPGVVVATSVITAENSSASIATIGGGPPQ 1825  
  
QY 48 QVNMDIPWNNYVOAQOGETFISPNTTPGDILFDLGLPHLPNFLSHLAOMYNWGNGNM 107  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 1826 QVDQOETWRNFYY---NDVFWSVADAPGSILTIVQHSPQNPFETAVLSOMYAGWAGM 1882  
  
QY 108 KVKVLLAGNAFTAGKIISIICPPIGFPAQAONISIA---QATMPHPHVIADVRLIEVPLE 164  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 1883 QRERIVAGSGVGFGRLVRVIPPG----IBPGGLEVRQPFPVHVIDARSLEPVTITPD 1937  
  
QY 165 VRNVLFH-NNDNPATMRVLCHMLYTPL--RASGSSSGTDFFVIAGRVLTCPSDFSFLV 221  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 1938 LRPNMYHTPDGPLVTLVSUYNLNINPFCGSTS-----AIQVTVEIRPEDFEVMIR 1992  
  
QY 222 PPNVQEOKTKPFSVNLPLNTLSNSRVPSLIKSMVMVSRDHGQMVFQ-----NGRVT 272  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 1993 AP3-SKTVDSSI PAGLLTT-----PVLTVGNDRNRNGQI VGLQPYPGGFCSTCRHN 2044  
  
QY 273 LD3LOQTTPPTSASOLCKIRGSVFHANGNGYNITEL-----DGSPYHAFESP----APIG 324  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 2045 LN3STYGWSSPRFGDI DHRRGSA SY-SGS NATNVLQFWYANAGS---AIDNPISQVAPDG 2100  
  
QY 325 FPDLBCECDWHMESPTTFQNTGDKVIKINYKQESAFAPHLGTIOAGLDLSVSWTNMTAK 384  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 2101 PPDMSFPFFNGPGIPAAGVGWFGAIWNSN-----SGAPNVTVTQA-----YE 2142  
  
QY 385 LGWVSPVSDGHRGDVDWPVIPRYGSTLTEAAQ-LAPP IYP PGFGEA 429  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 2143 LG?AT---GAPGNLQP-----TTWTSGAQTVAKSIIAYVTGTA 2177  
  
RESULT 6  
A53982  
capsid protein - European brown hare syndrome virus  
C:Species: European brown hare syndrome virus  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C;Accession: A53982  
J.R.Wirblisch, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Haas, B.; Thiel,  
R.; Virlich, C., 5164-5173, 1994  
A>Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A;Reference number: A53982; MUID:94309183; PMID:7518531  
A;Accession: A53982  
A>Status: preliminary  
A:Molecule type: mRNA  
A;Residues: 1-576 <WIR>  
A;Cross-references: GB:U09199  
C:Superfamily: human calicivirus capsid protein ;

Query Match 10.9%; Score 315; DB 2; Length 576;  
Best Local Similarity 26.2%; Pred. No. 1.7e-15;  
Matches 126; Conservative 69; Mismatches 200; Indels 86; Gaps 20;

QY 7 DAP-----TNMDGTSAGQLVPEANTAEPISMKEPVACGAATAATAGQVMNMDP----WTMN 58  
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 8 DAPCATATASPGTGTTDGMDFGVVASTDVVDTADNWAAVSATAGIGGPPQQASPQESRWVN 67  
  
QY 59 NYVQAPOGEEFTISPNTTPGDLFDLGLPHLPNFLSHLAOMYNWGNGMKVKVILLAGNAF 118  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 68 FFY---NDVFWSVTVDAPGSLTIVSQHS PQNPNFTQVLSQMYTAGWAGMGCFRIIVSGSI 124  
  
QY 119 TAGKIIIISCIPPFGFAAQ-NISIAQATMFPHVIADVRLIEVPLEDVRNVLFH--NNDNA 176

```

125 FGRLVCAIIPGQIQOPGLEVRQ---FPHVVIDARSLPVTITMDLREPWHHTGDPG 181
177 PTMLVCMVLTPLRASGSSGTDPF-----VIAGRVLTCPSDFSLFLVPPNVEOKTKP 231
182 LVPTLVSVVNNL-----INPFGTTSIAQVTVETRPSEDFEVLIRAFS--SKTVD 231
232 FSVPNPLNTLSNRVPSLIKSMVSRDHQOMVOFQ-----NGRVTLDGOLQGTTP 282
233 SVNPSMLLT-----PVLTGAGSDNRWAGAPVGLQVPVPGGFSTSNRNMNMNGSTYGWS 285
283 TSASOLCKIRGSVFHANGNGYNLTLDGSPYHAFESP---APIGFDPDLGECDDHMEAS 338
286 PREDDIDHSGNVSYPTGSAINTIETWYANAGTATTNPISNIAPDGFDDMGAIIPF-----S 341
339 PTTQFNTGDIKQINVKQESAFAPHLGTLTQADGLSDSVSNTNMIKLGWSPVSDGHRGD 398
342 GTT-ITPGAWGVGQVWNASNGTTPYGTQVQ-----YELGF---ANGAPSS 383
399 VDPWIPRYGSTLTERAAQL-APPYPPGGE-----AIVFMSDFPIAHGTNGLSVPCITP 453
384 IRP-----VTTTGAQLVAKSIYGVAIAQNOSSAGIIFLSKGMVS--TPGVAATTYP 434
454 Q 454
435 Q 435

RESULT 7
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
A:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

Query Match 10.8%; Score 311.5; DB 2; Length 2344;
Best Local Similarity 27.4%; Pred. No. 2.3e-14;
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;

QY 17 GAGQLVPEANTAEPISEMPVAGAAAT-----AAATAG-----QV 49
DB 1768 GKARTAPQGEAAGTATTASVPGTTDGLDPGVVATTSSVTTAENSSASIIATAGIGPPQV 1827
QY 50 NMIDPIMNNYVQAPQGEFTISPNNTPGDILDLQGLPHNLPFLSHLAOMYNGWGNMKV 109
DB 1828 DQOQETWRTNFY---NDVFTWSVADAPGSLYTVQHSPPNNPFTAVLSQMYAGWAGGMQF 1884
QY 110 KVLLAGNAFTAGKIITSCIPPGFAAQNISIA---QATMPPHVIADVRLVEPTEVPLEDVR 166
DB 1885 RFIVAGSGVGGRLVAAVTPPG-----IEIGPGLVQRQPHVVIDARSLPVTIIMPDLR 1939
QY 167 NVLFH-NNDNAPTMRLVCMLYTPL--RASGSSGTDPPFVIAGRVLTCPSDFSLFLVPP 223
DB 1940 PNMVHTGDPGLVPTLVLSYNNLINPFGSGTS-----AIQVTVETRPSEDFEVMIRAP 1994
QY 224 NVEQTKPSPVNPPLNTLSNRVPSLIKSMVSRDHQOMVOFQ-----NGRVTLD 274
DB 1995 S--SKTVDISPAAGLT-----PVLTVGNDNRWNGQIVGLQVPVPGGFSTCNRHNNLN 2046
QY 275 GLOQGTTPTSASOLCKIRGSVFHANGNGYNLTDEL-----DGSPYHAFESP-----APIGPP 326
DB 2047 GSTYGWSSPFAIDIHRRGSASYP-GSNATNVLQFWYANAGS----AVDNPISQVAPDGGP 2102
```

```

QY 327 DLGECDDHMEASPTQFNTGDIKQINVKQESAFAPHLGTLTQADGLSDSVSNTNMIKLG 386
DB 2103 DMSFVPFPGPGIPAAGWGWGCAIWNNSN-----SGAPNVTTVQA-----YELG 2144
QY 387 WVSVPVSDGHRGDVDPWVTPRYGSTLTERAAQ-LAPPIYPPGFGGEA 429
DB 2145 FAT-----GAPGNLQ-----TTNTSGAQTVAKSIYAVVTGTA 2177

RESULT 8
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S46944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus su
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 10.6%; Score 307.5; DB 2; Length 2344;
Best Local Similarity 26.3%; Pred. No. 4.5e-14;
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;

QY 12 MDG---TSAGQLVPEANTAE-----PTSMPEPVACAATAA-----ATAG----- 47
DB 1766 MEKARTAPQGEAAGTATTASVPGTTDMDPVGVVATTSSVTTAENSSASIIATAGIGPPQ 1825
QY 48 QVNMDPIMNNYVQAPQGEFTISPNNTPGDILDLQGLPHNLPFLSHLAOMYNGWGNM 107
DB 1826 QVDOQETWRTNFY---NDVFTWSVADAPGSLYTVQHSPPNNPFTAVLSQMYAGWAGM 1882
QY 108 KVKVLLAGNAFTAGKIITSCIPPGFAAQNISIA---QATMPPHVIADVRLVEPTEVPLED 164
DB 1883 QRFIVAGSGVGGRLVAAVTPPG-----IEIGPGLVQRQPHVVIDARSLPVTIIMP 1937
QY 165 VRNVLFH-NNDNAPTMRLVCMLYTPL--RASGSSGTDPPFVIAGRVLTCPSDFSLFLV 221
DB 1938 LRPNVHTGDPGLVPTLVLSYNNLINPFGSGTS-----AIQVTVETRPSEDFEVMIR 1992
QY 222 PPNVEOKTKPSPVNPPLNTLSNRVPSLIKSMVSRDHQOMVOFQ-----NGRVT 272
```

Db 1993 APS--SKTVDSISAPAGLLT-----PVLTVGVNDNRNGQIVGLQVPVPGGFSFNCNRHMN 2044  
Qy 273 LDGOLQGTTPTSASOLCKIRGVSFPHANGGYNLTTEL-----DGSYPYHAFESP-----APIG 324  
Db 2045 LNSTYGVSSPRFGDIGHRRGSASYP--GNNATNVLPQWYANAGS---AIDNPISQVAPDG 2100  
Qy 325 FPDLGECDDHMEASPTTQFNTGDVIOKQINVKQESAFAPHLGTIQADGLSDVSVNTNMIK 384  
Db 2101 FPDMSFVPFNGPGIIPAAAGVGFAGIWNNSN-----SGAPNVTTVQA-----YE 2142  
Qy 385 LGWSPVSDGHRGVDVDPWIPRYGSTLUEAAQ-LAPPIYPGGEAI-----VFMSDFPI 439  
Db 2143 LGFAT-----GAPGNLQP-----TTNTSGQTVAKSIYAVVTGTAQNPAGLFFVMASGVI 2191  
Qy 440 AHGNTGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPDTHRNLGFEFKLYPEGFTCV 499  
Db 2192 STPSANAITTPQPDRIVT-----TECTPAAAPVGVKNTPIMFASV 2232  
Qy 500 PNSS-----GTGQOTLPI 512  
Db 2233 RRTGDVNATAGSNGTQVGTGSQPLPV 2259  
  
RESULT 9  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: A48562  
A:Molecule type: genomic RNA  
A:Residues: 1-702 <NET>  
A:Cross-references: GB:H87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884  
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIIP:113565)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein  
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
Query Match 10.4%; Score 300.5; DB 1; Length 702;  
Best Local Similarity 25.2%; Pred. No. 2.6e-14;  
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;  
  
Qy 13 DGTSGAGQLVPEANTAEPIISMPEVAGAATAAGOVNMID-PW-IMNNYVQAPQGEFTI 70  
Db 154 DGPAGADIVTEQGTVOQOPVPAQSALTTLAAASTGKTVDCIEWTTFYSHTA-----VNW 209  
Qy 71 SPNPTPGDILFDLQGLPHNPLSHLAOMYNGWGNMKVKVLLAGNAFTAGKIISCITPP 130  
Db 210 STTEAQQKILFSRALSPENLPYRHSLSYSTWSGIDVRETVSGGVFGGKLAALIYPP 269  
Qy 131 GFQAONISTAQATMPHPVIAOVVRLEPTEVPLEDVRNLFHNDNAPTRMLVCMLYTPL- 189  
Db 270 GI--EPVESPTMLQPHVLFARQTEPVIFTPIDIRKTLYHSMDDTTRLRVIMVYNELI 327  
Qy 190 -----RAGSSSGTDPFVIAGRVLTCPSPDPSFLFLVPPNVEQTKPFSPVN--LPINT 241  
Db 328 NPYEQSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHSIPSDLI PRNS 377  
Qy 242 ---LSNSRVPSSLKSMVSRDHQMVQFQNGR-----VTLD----- 274  
Db 378 RHWNGNRWMTSTIDGFV-----QPRVQSNRHFDFDSTTGWSTPYPIEVTLEKLDORG 432  
Qy 275 GOLQGTTPTSASOL-----CKIRGSVFHANGGYNLTTEL-----DGSYPYHAFESPA 321  
Db 433 GQYFKVDTFEKSLVPLGCLDQGWDDTTIPTAWTASNGNDYTVAEYRITNNGTHFKGFYING 492  
Qy 322 PI-----GPPDLGECDDHMEASPTTQFNTGDVIOKQINVKQESAFAP-----HLGTIOA 369

Db 493 NLTTKVKGSDNLGET---QOTSRTLFASVG-----NYKDQNTINPHTHKITSNSLVVYDA 543  
Qy 370 DGLSDVSVNT-----NMIAKLGWV-----SPVSDGHRGVDVDPWIPRYGSTLTERAQA--P 419  
Db 544 NNYSAATAKTTTWHSTSMHSLGYLVDESVP-----GSDSTKVVRIATLP 587  
Qy 420 PIYPPGCGEAIVPEFMSDFPIAHGNTGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHY-LD 478  
Db 588 EAFNG--GNPPVFTTKIQIGH-----FDRHATTKCFNSQVLMTSOKLAENHYTLP 636  
Qy 479 PDT 481  
Db 637 PDS 639  
  
RESULT 10  
Qy 34 EPVAGAATAAGOVNMIDPWIMNNYVQAPQGEFTI-----SPNPTPGDILFDLQGLP 87  
Db 145 EPSAQMTAADMATGKSDVSEW-----EAFPSHTSVNWSSTETQCKILFKQSLGP 195  
Qy 88 HLNPFSLHQAOMYNGWGNMKVKVLLAGNAFTAGKIISCITPPGF-AAQNISTAQATWFP 146  
Db 196 LLNPLYEHLKSLKYVAMSGSVSEVREFSISGSGVFGGKLAALIYVPPGVDPVQSTMLQ----YP 252  
Qy 147 HVIADVVRLEPTEVPLEDVRNLFHNDNAPTRMLVCMLY-----TPLRAGSSSGTDPFV 202  
Db 253 HVLFARQVDPVIFSPDLRSTLHLPDPTDTSLSVIMVYNDLINPYANDSNSSGC---I 309  
Qy 203 IAGRVLTCPSPDPSFLFLVPP-----VT----- 223  
Db 310 VT--VETKPGDPDFHLLKPPGSMLTGHSVPDILPKSSSLWIGNRYMSDITDFVVRPFV 367  
Qy 224 -----NVEQTKPFSPVNL-PLN-TLSNSRVPSSLKSMVSRDHQMVQFQNGRVTL 274  
Db 368 FQANRHFNFQNETAGSAPRFRPRTITITISEKSKL--GIGVATDY-----IVP 414  
Qy 275 GOLQGTTPTSASOLCKIRGSVFHANGGYNLTDELGSPYHAFE-SPAPIGPPDLGEC- 332  
Db 415 GIPDGWPDPTTAEADLTAPAGD-YAITSNGNDIT--TGSEYDSTEVIRKNTNFRGMVICS 471  
Qy 333 ---WHMEASPTTQF-----NTGDVIOKQINVKQESAFAPHLGTIOADGLSD 374  
Db 472 LORAWGDKKISNTAFITTAIKEGNKIRPSNTIDMTK-LAVYQDT-----HVG-----EE 519  
Qy 375 VSVNTNMIKLGWSPVSDGHRGVDP-----WVTPRYGSTLTERAQAAPPYPPGFGAI 430  
Db 520 VQTSDDALLAGYTGIGEQAIGSDKDRVVRISVLPVEGAR-----GNHP 564  
Qy 431 VFMSDFPIAHGNTGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHY-LDPDT 481  
Db 565 IFYKNSIKLGVIRSIDV-----FNSQILHTSRQLSLNHYLLSPDS 605

## RESULT 11

VCWFF9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: genomic RNA  
A:Residues: 1-671 <CAR1>  
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266  
A:CROSS-references: PIDN:AAB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;  
Best Local Similarity 26.4%; Pred. NO. 1.6e-12;  
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;  
QY 13 DGTSGAGQLVPEANT-AEPTSMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70  
Db 127 DGSITA-----PEQGTWVGVIAPESQMSQSTADMATGKSVDSW-----EAFSFSH 173  
QY 71 -----SPNPTGDILOLQPHLPFLSLAQMYNGWGNMKVKVLLAGNAFTAGKIIII 125  
Db 174 TSNWSTSETQGIKLFKQSLGPLLPYLEHLAKLVAVWSGSIEVRSISGSGVFGGKLA 233  
QY 126 SCIPPCGF-AAONTISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLHNDNAPTMRVLCM 184  
Db 234 IVPPGVDPVQSTMLQ---YPHVLFDARQVEPVIFCLPDLKSTLYHVMSTDTTSVLIM 290  
QY 185 LY-----TPLRASGSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTFSPNPLN 240  
Db 291 VYNDLIPYANDANSNGC---IVT--VETKPGDFKFKHLLKPPG-----SMLTHG 335  
QY 241 TLNSRVPSLISKMYVSRDHQGVQFQNGRVTLGDLQGTTPTSASQLKIRGSVFHANG 300  
Db 336 SIPLDIPKTSLSWIGNRWSDITDF-----VIRPFVQANR 372  
QY 301 GNGYN-----LTLDGSPYHA-----FESPA-PIGFPDL---GEC--- 331  
Db 373 HFDENQETAGWSTPRFRISVTITQNGAKLGTGATDYPVCPDGPDPDTIPGELIPA 432  
QY 332 -DWHM-----EASPTQNTGCDVIK 350  
Db 433 GDVAITNGTNDITATGYDTADIIK 458

## RESULT 12

VCWFF9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: genomic RNA  
A:Residues: 1-671 <CAR1>  
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266  
A:CROSS-references: PIDN:AAB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

## VCWFF9

coat protein - feline calicivirus (strain Japanese F4)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: B40481  
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yam  
Virology 183, 810-814, 1991  
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
A:Reference number: A40481; MUID:91306470; PMID:1853578  
A:Accession: B40481  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <TOH>  
A:CROSS-references: GB:D90357; NID:g221264; PIDN:BAA14371.1; PID:g221266  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict  
Query Match 9.5%; Score 274; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. NO. 2.3e-12;  
Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;  
QY 23 PEANTA-EPISMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI-----SPNNT 75  
Db 133 PCGTAGVGVIAPESQMSQSTADMATGKSVDSW-----EAFSFTSVNWSSET 183  
QY 76 PGDILFDLOLQPHLPFLSLAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGF-AA 134  
Db 184 QGKILFKQSLGPLLPYLEHLKSLKLVAVWSGSIEVRSISGSGVFGGKLAIVVPPGVDPV 243  
QY 135 ONISTAQATMFPHVIADVRVLEPIEVPLEDVRNVLHNDNAPTMRVLCMLY----TPLR 190  
Db 244 QSTMLQ---YPHVLFDARQVEPVIFCLPDLKSTLYHVMSTDTTSVLIMVYNDLIPYA 300  
QY 191 ASGSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTFSPNPLNPLTSLNSRVPS 249  
Db 301 NDSNNGC---IVT--VETKPGDFKFKHLLKPPG-----SVLTHGSIIPS 339  
RESULT 13  
VCWFF9  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline cali  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NET>  
A:CROSS-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NET>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NET>  
A:CROSS-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict  
Query Match 9.3%; Score 269; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. NO. 5.4e-12;

Matches	74;	Conservative	38;	Mismatches	83;	Indels	44;	Gaps	9;
Qy	23	PEANT-ÆPISMEPVGAAGATAAAATAGQVNMIDPWIMNNYVQAPQCEFTI-----SPNNT	75						
Db	133	PEQGTWVGVIAPNAQMSAADMATCKSVDSW-----BAFFSFHTSVNWSSTSET	183						
Qy	76	PGDILFDLQGLPHLNPFLSHLAQNYGWNGVKNVYKVLLAGNAFTAGKIIISCIIPPGF-AA	134						
Db	184	QGKTLFKQSLGPLLNPYLTHLAKLYVAMSGVSDVRFESIGSGVFGGKLAALVWPPGIDPV	243						
Qy	135	QNIIAQAATMFPHVLAADVRLVETPEVLDVRNVLFHNNDNAPTMRVLCMLY-----TPLR	190						
Db	244	QSTSMQ---YPHVLFDFARQVPEVFISPDRLSTLYHLMSDSTDTSYIMVYNDLINPYA	300						
Qy	191	ASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVNPPLNTLSNSRVPFS	249						
Db	301	NDSNNSGC---IVT--VETKPGDPKFHLLKPPG-----SMLTGHSIFS	339						

```

RESULT :4
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562

```

R;Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel serins.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAI6220.1; PID:g334888  
A:Note: sequence extracted from NCBI backbone (NCBI:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89, 208, 329, 463, 482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	9.1%;	Score 264;	DB 1;	Length 703;
Best Local Similarity	27.4%;	Pred. No. 1.4e-11;		
Matches 104;	Conservative 53;	Mismatches 181;	Indels 42;	Gaps
Qy 13	DGTSAGQLVPEANTAEPIESMEPVAGATAATAAGOVNMIDPWIMNNVYVQAPQGEFTISP	72		
Db 154	DGPCSAELIVTEEGTVVOOQAPAPATLATLATASTGKSVGEQEWTF--SVHTSINNST	211		
Qy 73	NNTPGDILFDQLGPHLNNPFLSHLAQMYGNGVGNMKVKVLLAGNAFTAKGIISICPGF	132		
Db 212	VESQCKILYSQLNPINSINPYLDHIAKLYSTWSSGIDVRFVTSGSGVFGKLAALLVPPGV	271		
Qy 133	-AONISIAOATPEPHVADRVLEPIEVLIEDVRNLPHNNDNAPTMLVCMLYTPLRA	191		
Db 272	EPIESVMLQ---YPHVLFDARTEPVITPIDIKTLFHSMDETDTTKLVINPY----E	324		
Qy 192	SGSSSGDTPFVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFSPVN--LPPLNTL--SNSRV	247		
Db 325	NGVENKTTCSI---TVETRFSADFTALLKPPG--SLIKHGSIPSLIPRNSAHWMGNRW	379		
Qy 248	PSLIKSMWVRSDHGQVQFONGR-VILDGOLQG-TTPTTSASQLCKIRGSVFHANGNG--	303		
Db 380	WSTISGFSV-----QPRVFQSNRHFDFTSTTGWSTPYVPIEIKQGVK----GSNNKW	430		
Qy 304	YNLTLDGSPYHAPESPAPIGFFDGLCEDWHMEASTFT-QFNTGDVVK--QINVKQESAF	360		
Db 431	PHVIDTD-----KALVPGIPGWFDTTIPD---ETKATNGNFSYGESYRAGSTTIKPNENS	483		
Qy 361	APHLGCTIQAQGLSDVSNTN	380		
Db 484	THFKGTIVTCGLSTLSTVEIPEN	503		

```

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB1:09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match      9.0%  Score 261.5;  DB 2;  Length 668;
Best Local Similarity 31.1%;  Pred. No. 1.9e-11;
Matches 76;  Conservative 37;  Mismatches 86;  Indels 45;  Gaps

```

Qy	19	GQLV-PEANT-AEPI-SMEPVAGNAATAAATAGVNMDIPHMNNYVQAPOGGETI-----70
Dd	128	GSIVTPEQGTMVGVIATENPAQMSTAADMATCKSVDSWE-----EAFSEHTSVNW178
Qy	71	SPNPTGDILFDLQGLPHLPNFSLHIAOYMNGWGNMVKVLVLAGNAPTAKIIISCIPP130
Dd	179	STSETQCKILLFKOSLGPLNPYLEHLAKLYVAWSGIEVFRESISGSGVGKLAIVVP238
Qy	131	GF-AAQNISTAQATMPHFVTADVRVLETEVPLEEDRVNVLFHNNDONAPTRMLVCMLY---186
Dd	239	GVDPVQSTSMLO---YPHVLFDAQVEPFIIFPOLRSTLYHLMSSDOTTSLVWNAYNDL295
Qy	187	-TPURASGSSGGTDPPFVIAGRVLTCPSDFSFSLFLVPPPNEQKTFFSVPNLP.LNTLSNS245
Dd	296	INPYANDSNSGC---IVT--VETKCGSDFERHLLKPPG-----SVLTHG335
Qy	246	RVPS249
Dd	336	SVPS339

Search completed: March 10, 2003, 19:00:11  
Job time : 29 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:19:25 ; Search time 38 Seconds  
(without alignments)  
2955.149 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDAPTNDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.21.\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertebrate.\*
- 14: sp.unclassified.\*
- 15: sp.rvirus.\*
- 16: sp.bacteriaph.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12	Q8VA02 human calic
2	2846	98.3	545	12	Q9WI82 norwalk-lik
3	2846	98.3	545	12	Q9I8C8 norwalk-lik
4	2840	98.1	545	12	Q9I8C2 norwalk-lik
5	2798	96.6	545	12	Q9I8H12 human calic
6	2744	94.8	545	12	Q9IV45 human calic
7	2605	90.0	544	12	Q66418 desert shie
8	2124	73.3	539	12	Q9IV48 human calic
9	1979	68.3	544	12	Q9DU46 chiba virus
10	1975	68.2	544	12	Q9I185 human calic
11	1974	68.2	544	12	Q9QTE7 chiba virus
12	1972.5	68.1	540	12	Q9YS14 norwalk vir
13	1968	68.0	544	12	Q9I8C5 norwalk-lik
14	1964.5	67.8	546	12	Q9IV47 human calic
15	1956.5	67.6	543	12	Q9IV43 human calic
16	1956.5	67.6	544	12	Q8VA27 human calic

17	1954	67.5	544	12	Q9I8D1 norwalk-lik
18	1952	67.4	544	12	Q9I8J2 human calic
19	1946.5	67.2	543	12	Q9I8B9 norwalk-lik
20	1946	67.2	530	12	Q83884 norwalk vir
21	1942	67.1	544	12	Q9IV36 human calic
22	1941	67.0	544	12	Q9IV41 human calic
23	1938	66.9	530	12	Q9QV39 norwalk-lik
24	1930.5	66.7	531	12	Q8V735 norwalk-lik
25	1908.5	65.9	517	12	Q9JH41 norwalk-lik
26	1900	65.6	530	12	Q83876 norwalk vir
27	1886	65.1	541	12	Q9IV42 human calic
28	1274	44.0	522	12	Q8V629 norwalk-lik
29	1268	43.8	522	12	Q8V628 norwalk-lik
30	1260.5	43.5	519	12	Q9YQ22 bovine cali
31	1228	42.4	540	12	Q9I7X1 norwalk-lik
32	1223	42.2	542	12	Q66293 calicivirus
33	1222	42.2	542	12	Q9I3B6 human calic
34	1219	42.1	540	12	Q9IV50 human calic
35	1215.5	42.0	542	12	Q12305 snow mounta
36	1213	41.9	542	12	Q96877 norwalk-lik
37	1211	41.8	538	12	Q9I165 human calic
38	1203	41.5	540	12	Q9I7W8 norwalk-lik
39	1202.5	41.5	535	12	Q9I5C6 human calic
40	1202.5	41.5	540	12	Q9I8B0 norwalk-lik
41	1198.5	41.4	548	12	Q88291 small round
42	1196	41.3	537	12	Q9I3B7 human calic
43	1189	41.1	537	12	Q9I8H9 human calic
44	1188.5	41.0	548	12	Q9I7Z5 norwalk-lik
45	1188.5	41.0	548	12	Q9I6E5 human calic

## ALIGNMENTS

## RESULT 1

ID	Q8VA02	PRELIMINARY:	PRT;	545 AA.
AC	Q8VA02;			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Capsid protein.			
OS	Human calicivirus NLV/Potsdam 196/2000/DE.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norwalk-like viruses.			
OX	NCBI_TaxID=150135;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HU/NLV/POTSDAM 196/2000/DE;			
RA	Kuenkel U., Schreier E.;			
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF43267; AAL32455.1;			
DR	InterPro; IPR004005; Calici_coat.			
DR	Pfam; PF00915; Calici_coat; 1.			
DR	SEQUENCE 545 AA; 58601 MW; 7D47D9F3521D89 CRC64;			

Query Match	99.2%	Score 2873;	DB 12;	Length 545;
Best Local Similarity	99.1%	Pred. No. 1e-222;	2;	Indels 0;
Matches 540;	Conservative 3;	Mismatches 0;	Gaps 0;	
QY	1	MMASKDAPTNDGTSGAGOLVPEANTAEPI	SMPEVAGATAAATAGQVNMIDPWIMNNY	60
Db	1	MMASKDAPTNDGTSGAGOLVPEANTAEPI	SMPEVAGATAAATAGQVNMIDPWIMNNY	60
QY	61	VQAPQGEFTTSPNTPGDIILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA	120	
Db	61	VQAPQGEFTTSPNTPGDIILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA	120	
QY	121	GKIIISICPPGFAQNISIAQATMFPHVIA	DRVRLVLEPIEVPLEDRVRLVFNHNDNAPTMR	180
Db	121	GKIIISICPPGFAQNISIAQATMFPHVIA	DRVRLVLEPIEVPLEDRVRLVFNHNDNAPTMR	180

```

QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHQWQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Db 241 TLSNSRVPSLIKSMVSRDHQWQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKOESAF 360
Db 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKOESAF 360
QY 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPP 420
Db 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPP 420
QY 421 IYPGFGAIVFFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPGFGAIVFFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRNLGFEKLYPEGFMTCVPNSSGTGPTLPINGVFVSWVSREFYQLKPVGTAGPACRL 540
Db 481 THRNLGFEKLYPEGFMTCVPNSSGTGPTLPINGVFVSWVSREFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
Db 541 GIRRS 545

RESULT 2
Q9WI82
ID Q9WI82 PRELIMINARY; PRT; 545 AA.
AC Q9WI82;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/STAV/95/NOR;
RX MEDLINE=20349239; PubMed=10893150;
RA Myrmei M., Rimstad E.;
RT "Antigenic diversity of Norwalk-like viruses: expression of the capsid
protein of a genogroup I virus, distantly related to Norwalk virus.";
RL Arch. Virol. 145:711-723(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/STAV/95/NOR;
RA Myrmei M., Rimstad E.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF145709; AAD37377.1; .
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 58631 MW; 89D035D4FEB10BEF CRC64;

Query Match 98.3%; Score 2846; DB 12; Length 545;
Best Local Similarity 98.3%; Pred. No. 1.6e-220;
Matches 536; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMASKADPTNMDGTSGAGQLYPEANTAEPISEMPVAGAAATAAGQVNMIDPIMNNY 60
Db 1 MMASKADPTNMDGTSGAGQLYPEANTAEPISEMPVAGAAATAAGQVNMIDPIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSLHAQWNGWGMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSLHAQWNGWGMKVKVLLAGNAFTA 120
QY 121 GKIISCIIPGFAAQNISIAQATMPHPVIAADVRLPEIEVPLEDYRNVLFHNDNTPTMR 180

```

```

Db 121 GKIISCIIPGFAAQNISIAQATMPHPVIAADVRLPEIEVPLEDYRNVLFHNDNTPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHQWQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Db 241 TLSNSRVPSLIKSMVSRDHQWQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKOESAF 360
Db 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKOESAF 360
QY 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPP 420
Db 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPP 420
QY 421 IYPGFGAIVFFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPGFGAIVFFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRNLGFEKLYPEGFMTCVPNSSGTGPTLPINGVFVSWVSREFYQLKPVGTAGPACRL 540
Db 481 THRNLGFEKLYPEGFMTCVPNSSGTGPTLPINGVFVSWVSREFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
Db 541 GIRRS 545

RESULT 3
Q918C8
ID Q918C8 PRELIMINARY; PRT; 545 AA.
AC Q918C8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Honolulu/219/1992/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=171835;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/HONOLULU/219/1992/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/HONOLULU/219/1992/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/HONOLULU/219/1992/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF414403; AAL12962.1; .
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 58573 MW; 760A0DD2AB9E5AD7 CRC64;

Query Match 98.3%; Score 2846; DB 12; Length 545;
Best Local Similarity 98.5%; Pred. No. 1.6e-220;

```

Matches 537; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAATAAGQVNMIDPWIMNNY 60  
 Db 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAATAAGQVNMIDPWIMNNY 60

Qy 61 VQAPQGEFTTSPNNTGDIILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
 Db 61 VQAPQGEFTTSPNNTGDIILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120

Qy 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMR 180  
 Db 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPEFVITAGRVLTCPSPDFSLFLVPPNVEQKTKPFSPVNLPLN 240  
 Db 181 LVCMLYTPLRASGSSGTDPEFVITAGRVLTCPSPDFSLFLVPPNVEQKTKPFSPVNLPLN 240

Qy 241 TLSNSRVPSLIKSMWVRDHGMVQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300  
 Db 241 TLSNSRVPSLIKSMWVRDHGMVQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

Qy 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKQESAF 360  
 Db 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKQESAF 360

Qy 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPWIPRYGSTLTTEAAQLAPP 420  
 Db 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPWIPRYGSTLTTEAAQLAPP 420

Qy 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 Db 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

Qy 481 THRLNGEFLKYPEGFMTCVPNSSGTGPQTLPIGVFVSVSRFYQLKPVGTGAGPACRL 540  
 Db 481 THRLNGEFLKYPEGFMTCVPNSSGTGPQTLPIGVFVSVSRFYQLKPVGTGAGPACRL 540

Qy 541 GIRRS 545  
 Db 541 GIRRS 545

RESULT 4

Q918C2 PRELIMINARY; PRT; 545 AA.

AC Q918C2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Capsid protein.  
 CN ORF2.

OS Norwalk-like virus NLV/Little Rock/316/1994/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.  
 OX NCBI\_TaxID=171837;  
 RN [1]  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;  
 RX MEDLINE=97193806; PubMed=9041391;  
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577(1997).  
 RN [2]

RP SEQUENCE OF 100-192 FROM N.A.  
 RC STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;  
 RX MEDLINE=98071277; PubMed=9407386;  
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
 RA Seto Y., Monroe S.S., Glass R.I.;  
 RT "Correlation of patient immune responses with genetically  
 RT characterized small round-structured viruses involved in outbreaks of

RT nonbacterial acute gastroenteritis in the United States, 1990 to  
 RT 1995."; J. Med. Virol. 53:372-383(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;  
 RT "Genetic classification of 'Norwalk'-like viruses."; J. Infect. Dis. 181:S336-S348(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;  
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF144405; AAL12968.1; --;  
 DR InterPro: IPR004005; Calici\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 545 AA; 58666 MW; F55076D61D612E82 CRC64;

Query Match 98.1%; Score 2840; DB 12; Length 545;  
 Best Local Similarity 98.5%; Pred. No. 4.7e-220;  
 Matches 537; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAATAAGQVNMIDPWIMNNY 60  
 Db 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAATAAGQVNMIDPWIMNNY 60

Qy 61 VQAPQGEFTTSPNNTGDIILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
 Db 61 VQAPQGEFTTSPNNTGDIILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120

Qy 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMR 180  
 Db 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPEFVITAGRVLTCPSPDFSLFLVPPNVEQKTKPFSPVNLPLN 240  
 Db 181 LVCMLYTPLRASGSSGTDPEFVITAGRVLTCPSPDFSLFLVPPNVEQKTKPFSPVNLPLN 240

Qy 241 TLSNSRVPSLIKSMWVRDHGMVQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300  
 Db 241 TLSNSRVPSLIKSMWVRDHGMVQFONGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

Qy 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKQESAF 360  
 Db 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDHMEASPTTQFNTGDIKQINVKQESAF 360

Qy 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPWIPRYGSTLTTEAAQLAPP 420  
 Db 361 APHLGTTQADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPWIPRYGSTLTTEAAQLAPP 420

Qy 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 Db 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

Qy 481 THRLNGEFLKYPEGFMTCVPNSSGTGPQTLPIGVFVSVSRFYQLKPVGTGAGPACRL 540  
 Db 481 THRLNGEFLKYPEGFMTCVPNSSGTGPQTLPIGVFVSVSRFYQLKPVGTGAGPACRL 540

Qy 541 GIRRS 545  
 Db 541 GIRRS 545

RESULT 5

Q91HL2 PRELIMINARY; PRT; 545 AA.

AC Q91HL2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Capsid.

OS Human calicivirus NLV/VA98115/1998.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.  
 OX NCBI\_TaxID=165509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/VA98115/1998;  
 RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,  
 RA Fulton D., Morrow R., Matson D.O.;  
 RT "Baculovirus expression and antigenic characterization of the capsid  
 RT proteins of three Norwalk-like viruses."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY038598; AK84673.1; ...  
 DR InterPro: IPR004005; Calici\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 545 AA; 58295 MW; 258B0C7DEB34DAC CRC64;

Query Match 96.6%; Score 2798; DB 12; Length 545;  
 Best Local Similarity 96.7%; Pred. No. 1.le-216;  
 Matches 527; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MMASKDAPTNDGTSAGQLVPEANTAEPTISMEPVAGAATAAATAGOVNMDPWIMNNY 60  
 DB 1 MMASKDAPTNDGTSAGQLVPEANTAEPTISMEPVAGAATAAATAGOVNMDPWIMNDY 60  
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPELPHSLAQMNGWGMKVKVLLAGNAFTA 120  
 DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPELPHSLAQMNGWGMKVKVLLAGNAFTA 120  
 QY 121 GKIISCIIPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNNDNAPTMR 180  
 DB 121 GKIISCIIPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNNDNAPTMR 180  
 QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240  
 DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240  
 QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLTGQGLGTTPTSAQSLCKIRGVSFHANG 300  
 DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLTGQGLGTTPTSAQSLCKIRGVSFHANG 300  
 QY 301 GNGYNLTLDGSPYHAFESPAPIGPDIGECWHEASPTQFNTGDIKQINVKQESAF 360  
 DB 301 GNGYNLTLDGSPYHAFESPAPIGPDIGECWHEASPTQFNTGDIKQINVKQESAF 360  
 QY 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDVDPWVIPRYGSTLTEAAQLAPP 420  
 DB 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDVDPWVIPRYGSTLTEAAQLAPP 420  
 QY 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 DB 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 QY 481 THRLNGEFLKYPGFMTCVPSNSGTCPTLPINGVFVSVWSRYFQYKLPVGTAGPACRL 540  
 DB 481 THRLNGEFLKYPGFMTCVPSNSGTCPTLPINGVFVSVWSRYFQYKLPVGTAGPACRL 540  
 QY 541 GIRRS 545  
 DB 541 GIRRS 545

RESULT 6  
 Q91V45 PRELIMINARY; PRT; 545 AA.  
 ID Q91V45  
 AC Q91V45;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/VA98115/1998.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.

OX NCBI\_TaxID=122916;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HU/NLV/BIRMINGHAM/93/UK;  
 RX MEDLINE=20404883; PubMed=10949950;  
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
 RT "Capsid protein diversity among 'Norwalk-like' viruses."  
 RL Virus Genes 20:227-236(2000).  
 DR EMBL: AJ277612; CAB89093.1; ...  
 DR InterPro: IPR004005; Calici\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 545 AA; 58689 MW; DC6EBC6EBECDF7C7 CRC64;

Query Match 94.8%; Score 2744; DB 12; Length 545;  
 Best Local Similarity 95.8%; Pred. No. 2.5e-212;  
 Matches 522; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 MMASKDAPTNDGTSAGQLVPEANTAEPTISMEPVAGAATAAATAGOVNMDPWIMNNY 60  
 DB 1 MMASKDAPTNDGTSAGQLVPEANTAEPTISMEPVAGAATAAATAGOVNMDPWIMNNY 60  
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPELPHSLAQMNGWGMKVKVLLAGNAFTA 120  
 DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPELPHSLAQMNGWGMKVKVLLAGNAFTA 120  
 QY 121 GKIISCIIPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNNDNAPTMR 180  
 DB 121 GKIISCIIPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNNDNAPTMR 180  
 QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240  
 DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240  
 QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLTGQGLGTTPTSAQSLCKIRGVSFHANG 300  
 DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLTGQGLGTTPTSAQSLCKIRGVSFHANG 300  
 QY 301 GNGYNLTLDGSPYHAFESPAPIGPDIGECWHEASPTQFNTGDIKQINVKQESAF 360  
 DB 301 GNGYNLTLDGSPYHAFESPAPIGPDIGECWHEASPTQFNTGDIKQINVKQESAF 360  
 QY 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDVDPWVIPRYGSTLTEAAQLAPP 420  
 DB 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDVDPWVIPRYGSTLTEAAQLAPP 420  
 QY 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 DB 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
 QY 481 THRLNGEFLKYPGFMTCVPSNSGTCPTLPINGVFVSVWSRYFQYKLPVGTAGPACRL 540  
 DB 481 THRLNGEFLKYPGFMTCVPSNSGTCPTLPINGVFVSVWSRYFQYKLPVGTAGPACRL 540  
 QY 541 GIRRS 545  
 DB 541 GIRRS 545

RESULT 7  
 Q66418 PRELIMINARY; PRT; 544 AA.  
 ID Q66418  
 AC Q66418;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Capsid protein.  
 OS Desert Shield virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.  
 OX NCBI\_TaxID=33755;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```
RC STRAIN_DSV395;
RA MEDLINE=94174735; PubMed=8128635;
RX Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
RT "Molecular characterization and expression of the capsid protein of a
RT Norwalk-like virus recovered from a Desert Shield troop with
RT gastroenteritis.";
RL Virology 200:319-325(1994).
DR EMBL: U0469; AAA16285.1; -.
DR InterPro: IPR004005; Calici.coat.
DR Pfam: PF00915; Calici.coat; 1.
SQ SEQUENCE 544 AA; 58606 MW; 07A4AE7E4BCE81E8 CRC64;

Query Match 90.0%; Score 2605.5; DB 12; Length 544;
Best Local Similarity 89.2%; Pred. No. 3.5e-201;
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGAATAAATAGOVNMDPWTMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGAATAAATAGOVNMDPWTMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVRLFNHNDNAPTMR 180
DB 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVRLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDIKQINVKQESAF 360
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHMEASPTTQFNTGDIKQINVKQESAF 360
QY 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLAPP 420
DB 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLAPP 420
QY 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 420 IYPGFGGEAIVFFMSDFPVVSGVNGMRIPCTLPQEQYVAHFVNEQAPTRGEAALLHYVDPD 479
DB 420 IYPGFGGEAIVFFMSDFPVVSGVNGMRIPCTLPQEQYVAHFVNEQAPTRGEAALLHYVDPD 479
QY 461 THRNLGKFLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPACRL 540
DB 461 THRNLGKFLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 540 GIRRS 544

RESULT 8
Q91V48 PRELIMINARY; PRT; 539 AA.
AC Q91V48;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Winchester/94/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/WINCHESTER/94/UK;
RX MEDLINE=20404883; PubMed=10949950;
```

```
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277609; CAB89090.1; -.
DR InterPro: IPR004005; Calici.coat.
DR Pfam: PF00915; Calici.coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 1.9e-162;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGAATAAATAGOVNMDPWTMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGAATAAATAGOVNMDPWTMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVRLFNHNDNAPTMR 180
DB 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVRLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHM--EASPTTQFNTGDIKQINVKQES 358
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHM--EASPTTQFNTGDIKQINVKQES 358
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHM--EASPTTQFNTGDIKQINVKQES 358
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDLGECDWHM--EASPTTQFNTGDIKQINVKQES 358
QY 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLA 418
DB 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLA 418
QY 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLA 418
DB 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSLTTEAAQLA 418
QY 419 PIYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 478
DB 419 PIYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 478
QY 415 PLIFPGFGGEAIVFFMSDFPIVSG-NTAQIPCTLPQEFVSHFVEQQAIRGEAALLHYVD 473
DB 415 PLIFPGFGGEAIVFFMSDFPIVSG-NTAQIPCTLPQEFVSHFVEQQAIRGEAALLHYVD 473
QY 479 PDTHRNLGKFLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPAC 538
DB 479 PDTHRNLGKFLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPAC 538
QY 474 PDTHRNLGKFLYPEGFMTCVPNTGG-GPQNLPTNGVVFVSWSRFYQLKPVGTAGPAR 532
DB 474 PDTHRNLGKFLYPEGFMTCVPNTGG-GPQNLPTNGVVFVSWSRFYQLKPVGTAGPAR 532
QY 539 RL3IRR 544
DB 533 RL3VRR 538

RESULT 9
Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RA Someya T., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DDAJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

RC STRAIN=HU/NLV/CHIBA 407/1987/JP;  
 RX MEDLINE=20569531; PubMed=11118371;  
 RA Someya Y., Takeda N., Miyamura T.;  
 RT "Complete nucleotide sequence of the Chiba virus genome and functional  
 expression of the 3C-like protease in *Escherichia coli*.";  
 RL Virology 278:490-500(2000).  
 DR EMBL; AB042808; BAB18267.1; -.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 68.3%; Score 1979; DB 12; Length 544;  
 Best Local Similarity 66.8%; Pred. No. 9e-151;  
 Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNNIDPWIMNNY 60  
 Db 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNNIDPWIMNNY 60

Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120  
 Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120

Qy 121 GKIIISCIIPPFAAQNISIAQATMFPHVIADRVRLIEVPLEDVRNLFHNNNDAPTMR 180  
 Db 121 GKIIISCIIPPFAAQNISIAQATMFPHVIADRVRLIEVPLEDVRNLFHNNNDAPTMR 180

Qy 181 LVCMLYTPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
 Db 181 LVCMLYTPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240

Qy 241 TLSNRVPSLIKSMVSRDHGQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300  
 Db 241 TLSNRVPSLIKSMVSRDHGQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300

Qy 301 GNGYNLTDELGSYPYHAFESPAPIGFPDLGECQWHEAS--PTQFNTGDVIKQINVKQES 358  
 Db 301 GNGYNLTDELGSYPYHAFESPAPIGFPDLGECQWHEAS--PTQFNTGDVIKQINVKQES 358

Qy 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGVDVDPWIPRYGSTLTAAQLA 418  
 Db 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGVDVDPWIPRYGSTLTAAQLA 418

Qy 419 PPIYPGFGGEAIVFMSDFPIAHGTNGLS---VPCTTPOEFTVHFVNEQAPTRGEAALL 474  
 Db 419 PPIYPGFGGEAIVFMSDFPIAHGTNGLS---VPCTTPOEFTVHFVNEQAPTRGEAALL 474

Qy 475 HYLDPDTHRLNGEFKLYPEGFMTCVPNSSGTPQTLPIGVFVFSWVSRYQLKPVGTA 534  
 Db 475 HYLDPDTHRLNGEFKLYPEGFMTCVPNSSGTPQTLPIGVFVFSWVSRYQLKPVGTA 534

Qy 535 GPA-CRLGIRR 544  
 Db 535 GPA-CRLGIRR 544

RESULT 10  
 Q91185 ID Q91185 PRELIMINARY; PRT; 544 AA.  
 AC Q91185;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus HU/NLV/Koblentz/433/2000/DE.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.  
 OC NCBI\_TaxID=165252;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HU/NLV/KOBLENZ 433/2000/DE;  
 RA Kuenkel U., Schreier E.;  
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated

RT with Norwalk-like viruses in Germany.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF394960; AAK72048.1; -.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 68.2%; Score 1975; DB 12; Length 544;  
 Best Local Similarity 66.8%; Pred. No. 1.9e-150;  
 Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNNIDPWIMNNY 60  
 Db 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNNIDPWIMNNY 60

Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120  
 Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120

Qy 121 GKIIISCIIPPFAAQNISIAQATMFPHVIADRVRLIEVPLEDVRNLFHNNNDAPTMR 180  
 Db 121 GKIIISCIIPPFAAQNISIAQATMFPHVIADRVRLIEVPLEDVRNLFHNNNDAPTMR 180

Qy 181 LVCMLYTPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
 Db 181 LVCMLYTPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240

Qy 241 TLSNRVPSLIKSMVSRDHGQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300  
 Db 241 TLSNRVPSLIKSMVSRDHGQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300

Qy 301 GNGYNLTDELGSYPYHAFESPAPIGFPDLGECQWHEAS--PTQFNTGDVIKQINVKQES 358  
 Db 301 GNGYNLTDELGSYPYHAFESPAPIGFPDLGECQWHEAS--PTQFNTGDVIKQINVKQES 358

Qy 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGVDVDPWIPRYGSTLTAAQLA 418  
 Db 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGVDVDPWIPRYGSTLTAAQLA 418

Qy 419 PPIYPGFGGEAIVFMSDFPIAHGTNGLS---VPCTTPOEFTVHFVNEQAPTRGEAALL 474  
 Db 419 PPIYPGFGGEAIVFMSDFPIAHGTNGLS---VPCTTPOEFTVHFVNEQAPTRGEAALL 474

Qy 475 HYLDPDTHRLNGEFKLYPEGFMTCVPNSSGTPQTLPIGVFVFSWVSRYQLKPVGTA 534  
 Db 475 HYLDPDTHRLNGEFKLYPEGFMTCVPNSSGTPQTLPIGVFVFSWVSRYQLKPVGTA 534

Qy 535 GPA-CRLGIRR 544  
 Db 535 GPA-CRLGIRR 544

RESULT 11  
 Q9QTE7 ID Q9QTE7 PRELIMINARY; PRT; 544 AA.  
 AC Q9QTE7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Capsid protein.  
 OS Chiba virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norwalk-like viruses.  
 OC NCBI\_TaxID=99565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CHIBA 407;  
 RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,  
 RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;  
 RT "Expression and self-assembly of capsid proteins of the Chiba virus, a  
 genetically distinct Norwalk-like virus.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022679; BAA82106.1; -.

```
DR InterPro: IPR004005; Calici_coat.  
DR Pfam: PF00915; Calici_coat; 1.  
SQ SEQUENCE 544 AA; 58323 MW; ABD1C1FC4F93D872 CRC64;  
  
Query Match 68.28; Score 1974; DB 12; Length 544;  
Best Local Similarity 66.68; Pred. No. 2.3e-150;  
Matches 367; Conservative 68; Mismatches 102; Indels 14; Gaps 6;  
  
Qy 1 MMASKDPTNMDDTSCAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDIPWIMNNY 60  
Db 1 MMASKDPTNMDDTSCAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDIPWIMNNY 60  
Qy 61 VQAPQGETTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
Db 61 VQAPQGETTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
Qy 121 GKIIISCPGFAAQNTSIAQATMFPHVIAADVRLVLEPIEVLDPVLEDRVNLVFNHNDNAPTMR 180  
Db 121 GKIIISCPGFAAQNTSIAQATMFPHVIAADVRLVLEPIEVLDPVLEDRVNLVFNHNDNAPTMR 180  
Qy 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLVPPNVVEQKTRKPSVNPPLN 240  
Db 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLVPPNVVEQKTRKPSVNPPLN 240  
Qy 241 TILSNRVPSLIKSMVSRDHQMVQFNGRVTLTGQLOGTTPTSASOLCKIRGVSFVHANG 300  
Db 241 TILSNRVPSLIKSMVSRDHQMVQFNGRVTLTGQLOGTTPTSASOLCKIRGVSFVHANG 300  
Qy 301 GNGVNLTELDGSPYHAFESPAIGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 358  
Db 301 GNGVNLTELDGSPYHAFESPAIGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 358  
Qy 358 QRVNLNLTDELDSGPFMGFCAPAGAPAGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 358  
Db 358 QRVNLNLTDELDSGPFMGFCAPAGAPAGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 358  
Qy 418 AFAPHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGDDVPMVPIRGYSTLFEAAQLA 418  
Db 418 AFAPHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGDDVPMVPIRGYSTLFEAAQLA 418  
Qy 474 PPIYPGGEAIVFFMSDFPIAHGTNGLS-----VPCTIPQEFVTHFVNEQAPTGEAALL 474  
Db 474 PPIYPGGEAIVFFMSDFPIAHGTNGLS-----VPCTIPQEFVTHFVNEQAPTGEAALL 474  
Qy 534 HYLPDPTNRNLGELFKLYPEGFTMTCVPNSSGTCGPTLPINGVVFVSVWSRYQLKPVGTA 534  
Db 534 HYLPDPTNRNLGELFKLYPEGFTMTCVPNSSGTCGPTLPINGVVFVSVWSRYQLKPVGTA 534  
Qy 535 GPA-CRLGIRR 544  
Db 535 GPA-CRLGIRR 544  
  
RESULT 12  
Q9YS14 PRELIMINARY; PRT; 540 AA.  
AC Q9YS14  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Capsid protein.  
OS Norwalk virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.  
OX NCBI_TaxID=11983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NORWALK-Like;  
RA Schreier E., Doering F., Kuenkel U.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
with small round structured viruses in Germany in 1997/98.";  
RL Arch. Virol. 145:443-453(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NORWALK-Like;  
RA Schreier E., Doering F., Kuenkel U.;  
  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF093797; AAC64603.1; --  
DR InterPro: IPR004005; Calici_coat.  
DR Pfam: PF00915; Calici_coat; 1.  
SQ SEQUENCE 540 AA; 58018 MW; 647843487A654CBE CRC64;  
  
Query Match 68.18; Score 1972.5; DB 12; Length 540;  
Best Local Similarity 67.78; Pred. No. 3e-150;  
Matches 377; Conservative 60; Mismatches 89; Indels 31; Gaps 7;  
  
Qy 1 MMASKDPTNMDDTSCAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDIPWIMNNY 60  
Db 1 MMASKDPTNMDDTSCAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDIPWIMNNY 60  
Qy 61 VQAPQGETTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
Db 61 VQAPQGETTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120  
Qy 121 GKIIISCPGFAAQNTSIAQATMFPHVIAADVRLVLEPIEVLDPVLEDRVNLVFNHNDNAPTMR 179  
Db 121 GKIIISCPGFAAQNTSIAQATMFPHVIAADVRLVLEPIEVLDPVLEDRVNLVFNHNDNAPTMR 179  
Qy 180 RLYCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLVPPNVVEQKTRKPSVNPPLN 239  
Db 180 RLYCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLVPPNVVEQKTRKPSVNPPLN 239  
Qy 240 NTLSNRVPSLIKSMVSRDHQMVQFNGRVTLTGQLOGTTPTSASOLCKIRGVSFVHANG 299  
Db 240 NTLSNRVPSLIKSMVSRDHQMVQFNGRVTLTGQLOGTTPTSASOLCKIRGVSFVHANG 299  
Qy 298 GNGVNLTELDGSPYHAFESPAIGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 350  
Db 298 GNGVNLTELDGSPYHAFESPAIGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 350  
Qy 350 QRVNLNLTDELDSGPFMGFCAPAGAPAGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 350  
Db 350 QRVNLNLTDELDSGPFMGFCAPAGAPAGFPGDLGECDMHWEAS--PTTQNTGVDVTKQINVKQES 350  
Qy 351 QINVKQESAFAPHLGTI---QADGLSDVSVNTNMIAGLWSPVSDGHRGDDVPMVPIRGYSTLFEAAQLA 407  
Db 351 QINVKQESAFAPHLGTI---QADGLSDVSVNTNMIAGLWSPVSDGHRGDDVPMVPIRGYSTLFEAAQLA 407  
Qy 407 GSTLFEAAQLAPPIYPGGEAIVFFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPT 467  
Db 407 GSTLFEAAQLAPPIYPGGEAIVFFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPT 467  
Qy 467 RGEAALLHYLDPDPTNRNLGELFKLYPEGFTMTCVPNSSGTCGPTLPINGVVFVSVWSRYQLKPVGTA 527  
Db 467 RGEAALLHYLDPDPTNRNLGELFKLYPEGFTMTCVPNSSGTCGPTLPINGVVFVSVWSRYQLKPVGTA 527  
Qy 527 GPA-CRLGIRR 544  
Db 527 GPA-CRLGIRR 544  
  
RESULT 13  
Q918C5 PRELIMINARY; PRT; 544 AA.  
AC Q918C5  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Capsid protein.  
OS Norwalk-like virus NLV/Baltimore/277/1993/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.  
OX NCBI_TaxID=171836;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=HU/NLV/BALTIMORE/277/1993/US;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).
```

```
RN RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=HU/NLV/BALTIMORE/277/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF414404; AAL12365.1;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58208 MW; D5170FA4B958A672 CRC64;

Query Match 68.0%; Score 1968; DB 12; Length 544;
Best Local Similarity 66.4%; Pred. No. 6.9e-150;
Matches 366; Conservative 69; Mismatches 102; Indels 14; Gaps 6;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTSADGATGAGQLVPEVNTADPIPIPVAGSSTALATAGRVNLDPWIIINF 60
QY 61 VQAPQGEFTSPNNTPGDILFDLQGLPHLNPFLSHLAQWNGVGNKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTSPNNTPGDILFDLQGLPHLNPFLSHLAQWNGVGNKVKVLLAGNAFTA 120
QY 121 GKIIISCIPIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDRVNLFFHNNDNAPTMR 180
DB 121 GKVIICCVPPGFSQRTLSIAQATLFPHVIADRVLEPIEVPLEDRVNLFFHNNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDSEFVAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPLN 240
DB 181 LLCLMLYTPLRASGSSGTDSEFVAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPLK 240
QY 241 TLSNSRVPSLIKMVMVSRDHGMQVQFQNGRVTLTGQGLTPTTSASOLCKIRGSVFHANG 300
DB 241 CLNSRIPNPIEGNLSLSDQTNQVQFQNGRCTLIDGQPLGTPVSVSOLCKFRGRI--TSG 298
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGCDWHMEAS--PTQFNTGDVQIKQINVKQES 358
DB 299 QGVNLTELGDSPFMAFAAPAGAPFDLGSCDWHIEMSKIPNSSTQNNPLVNSVKPNQ 358
QY 359 AFAPHLGTIQADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVPRYGSTLTAEAAQLA 418
DB 359 QFVPHLSLITLD--ENVSSGGDYGITQWTSPPSDSGGANTNPKIPDYGSSLAESAQLA 416
QY 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS----VPCITIQEFVTHFVNQCAPTRGEAALL 474
DB 417 PAVYPGFGFNEVIVFMASIP---GPNQSGSPNLVPCLLPQEYITTHFISQAPIQGEAALL 473
QY 475 HYLDPDTHRNLFGEFKLYPEGMFTCPVNSSGTGPTLPINGVFVFSWSRFRYOLKPVGTA 534
DB 474 HYVDPDTHRNLFGEFKLYPGGILTCVNSSTSGTGQQLPLDGVFVFSWSRFRYOLKPVGTA 533
QY 535 GPA-CRIGIRR 544
DB 534 GPARGRLGVR 544
```

```
RESULT 14
Q9IV47 PRELIMINARY; PRT: 546 AA.
AC Q9IV47;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Whiterose/96/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122914;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/WHITEROSE/96/UK;
RX MEDLINE=20404883; PubMed=1094950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277610; CAB89091.1;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58737 MW; 501F9D2BE2A77B51 CRC64;

Query Match 67.8%; Score 1964.5; DB 12; Length 546;
Best Local Similarity 66.2%; Pred. No. 1.3e-149;
Matches 365; Conservative 78; Mismatches 95; Indels 13; Gaps 7;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKGAPQSGADSGAGQLVPEVNTADPLPMEPVAGTAVATAGQVNMIDPWVNNF 60
QY 61 VQAPQGEFTSPNNTPGDILFDLQGLPHLNPFLSHLAQWNGVGNKVKVLLAGNAFTA 120
DB 61 VQSPQGEFTSPNNTPGDILFDLQGLPHLNPFLSHLSQWNGVGNKVKVLLAGNAFTA 120
QY 121 GKIIISCIPIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDRVNLFFHNNDNAPTMR 180
DB 121 GKIIICCVPPGFTSSSLTIAQATLFPHVIADRVLEPIEVPLEDRVNLFFHNNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDSEFVAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDSEFVAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPLQ 240
QY 241 TLSNSRVPSLIKMVMVSRDHGMQVQFQNGRVTLTGQGLTPTTSASOLCKIRGSVFHANG 300
DB 241 TLSNSRFPSLIOGMLSPDASQVQFQNGRCLIDGQLGTTPTTSQGLFRVGRKI--NQG 298
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGCDWHMEASPT-TQFNTGDVQIKQINVKQES 358
DB 299 ARTNLNTEVDGKPFMAFDSAPVGFDFGKCDWHMRVSKTPNNTSSGDMRVSQVQTNVQ 358
QY 359 AFAPHLGTIQADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVPRYGSTLTAEAAQLA 418
DB 359 GFVPHLSIQDFEVENHPTG-DYIGTIEWISQSPPLGTDINLWEIPDYGSSLSQAANLA 417
QY 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS----VPCITIQEFVTHFVNQCAPTRGEAALL 474
DB 418 PVFPFGGEALVIVFSAPP---GPNNSAPNDVPCLLPQEYITTHFVSEQAPTMDAALL 474
QY 475 HYLDPDTHRNLFGEFKLYPEGMFTCPVNSSGTGPTLPINGVFVFSWSRFRYOLKPVGTA 534
DB 475 HYVDPDTHRNLFGEFKLYPGGILTCVNGVAGAPQQLPLNGVFLFVFSWSRFRYOLKPVGTA 534
QY 535 GPA-CRIGIRR 544
DB 535 STARGRLGVR 545

RESULT 15
Q9IV43 PRELIMINARY; PRT: 543 AA.
ID Q9IV43
```



Search completed: March 10, 2003, 18:27:37  
Job time : 40 secs



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1974.5	68.2	546	1	COAT_SOUV3		Q04542 southampton v
2	1148.5	39.7	539	1	COAT_LORDV		P54635 lordsdale v
3	316	10.9	2344	1	POLN_RBDV		P27410 rabbit hemo
4	300.5	10.4	702	1	COAT_MSXV1		P62884 san miguel
5	289	10.0	2208	1	POLN_MNSCV		Q69014 manchester
6	276	9.5	671	1	COAT_FCVCV9		P27405 feline cali
7	274	9.5	668	1	COAT_FCVCV4		P27405 feline cali
8	269	9.3	668	1	COAT_FCVC6		P27404 feline cali
9	264	9.1	703	1	COAT_SMSV4		P62885 san miguel
10	146.5	5.1	2205	1	POLG_POL2W		P23069 poliovirus
11	145.5	5.0	2207	1	POLG_POL2L		P06210 poliovirus
12	145	5.0	2206	1	POLG_POL3L		P03302 poliovirus
13	143.5	5.0	2194	1	POLG_HE701		P32537 human enter
14	139.5	4.8	2206	1	POLG_POL32		P06209 poliovirus
15	137.5	4.7	2209	1	POLG_POLL5		P03301 poliovirus
16	136.5	4.7	2208	1	POLH_POLLM		P03300 p genome po
17	135.5	4.7	2206	1	POLG_POLIM		P03299 p genome po
18	133.5	4.6	2179	1	POLG_HRV14		P03303 human rhino
19	124.5	4.3	855	1	POLG_HRV3		Q82081 human rhino
20	124	4.3	2214	1	POLG_CXA24		P36290 c genome po
21	123.5	4.3	2175	1	POLG_BOVEV		P12915 bovine ente
22	119.5	4.1	2196	1	POLG_EC05N		Q9Y1J1 e genome po
23	118.5	4.1	2194	1	POLG_EC30B		Q9WN78 e genome po
24	117.5	4.1	2185	1	POLG_CXB3P		Q03053 c genome po
25	117	4.0	788	1	BGSB_XANAC		P58933 xanthomonas
26	117	4.0	2183	1	POLG_CXB4E		Q86887 c genome po
27	117	4.0	2183	1	POLG_CXB4J		P08292 c genome po
28	116.5	4.0	2206	1	POLG_CXA21		P22055 c genome po
29	115	4.0	2185	1	POLG_SVDVH		P16604 s genome po
30	115	4.0	2185	1	POLG_SVDVU		P13900 s genome po
31	113.5	3.9	2164	1	POLG_HRV89		P07210 human rhino
32	112.5	3.9	2184	1	POLG_EC01F		Q91734 e genome po
33	112	3.9	2195	1	POLG_EC01G		P29813 e genome po

```
QY 181 LVCMLYTLPRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVQKTKPFSVPNPLN 240
DB 181 LVCMLYTLPRASGSSGSDSFWAGRVLTAPSSDFSEFLVPPNIEQKTRAFVNIPLQ 240
QY 241 TLSNRVSLIKSMVSRDHGMQVQFNGRVTLQGLQGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVSLIKSMVSRDHGMQVQFNGRVTLQGLQGTTPTSASOLCKIRGSVFHANG 300
QY 301 GNGVNLTELDGSPYHAFESPAPGPDGLGECDHMEASPT--TQNTGDIKQINVKOE-S 358
DB 301 GNGVNLTELDGSPYHAFESPAPGPDGLGECDHMEASPT--TQNTGDIKQINVKOE-S 358
QY 359 AFAPHLGTIQADGLSDSVSNNTMIKLGWSPVSDGHRGDVDVPIPRYQSTLTEAAQLA 418
DB 359 AFAPHLGTIQADGLSDSVSNNTMIKLGWSPVSDGHRGDVDVPIPRYQSTLTEAAQLA 418
QY 419 PPIYPGGEAIVFMSDFPIAHGTNGLS-----VPCITPOEFVTHFNEQAPTRGEAALL 474
DB 419 PPIYPGGEAIVFMSDFPIAHGTNGLS-----VPCITPOEFVTHFNEQAPTRGEAALL 474
QY 475 HYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTLPINGVFVSVWSRYQLKPVGTA 534
DB 475 HYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTLPINGVFVSVWSRYQLKPVGTA 534
QY 535 GPA-CRLGIRR 544
DB 535 STARKLGVR 545

RESULT 2
COAT_LORDV STANDARD; PRT; 539 AA.
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capsid protein).
OS Lordsdale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and
RT expression of virus-like particles from a genetic group II small
RT round structured virus."
RL J. Gen. Virol. 76:2349-2355(1995).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86557; CAA60255.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
```

Query Match 39.7%; Score 1148.5; DB 1; Length 539;  
Best Local Similarity 44.9%; Pred. No. 2.8e-80;  
Matches 254; Conservative 77; Mismatches 186; Indels 49; Gaps 16;

```
QY 1 MMASKADPTNMDGTSGAGQLVPEANTAPPIISMPEVAGAATAAATAGQVNMIDPWIMNNY 60
DB 1 MKMASDANPS--DGS--AANLVPEVNN--EVMALEPVVGAIAAPVAGQNVIDPWIRNF 56
QY 61 VQAPQCEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMNGVGNMVKVLLAGNAFTA 120
DB 61 VQAPGGEFTVSPNAPGEITLWSAPLGPDLNPYLSHLSRMNGYAGGFEVQVILAGNAFTA 116
QY 121 GKIIISICPGPAAQNSTIAQATMFPHVIADVRLBPIEVPLEDEVNVLFFH--NNDNAPTM 179
DB 117 GKVIFAAVPNPFTGELSPSOVTFMFPHIIVDVRLBPIEVPLEDEVNVLFFH--NNDNAPTM 176
QY 180 RLVCMYLTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVQKTKPFSVPNPL 239
DB 177 KLAMLYTTLRA--NNAGDDVFTVSCRVLTRSPDPDFLFLVPTVESRTKPTVPVLTV 234
QY 240 NTLNSRVSLIKSMVSRDHGMQVQFNGRVTLQGLQGTTPTSASOLCKIRGSVFHANG 299
DB 235 EEMSNSRFPDLEKLTGPFSSAFVQVQNGRCITDGLVLTGTTQLSVAVNICNFRGDVTHIA 294
QY 300 GNGV--NLTELDGSPYHAFES--PAPIGPDL-----GECDHMEASPTTQNT 345
DB 295 GSHDYTMNLASQNSWVDPTEETPAPLGPDPFGVKIQLGLTQTTRADGSTRAHKAT-VST 353
QY 346 GDVIKQINVKQSAFAPHLGTIQ--ADGLSDSVSNNTMIKLGWSPVSDG--HRGDVDP 401
DB 354 GSV-----HETPKLGSVQFTTDTNDFQAGN--TKETPVGVIGDGHQNEPQQ 401
QY 402 WTIPTYGSTLTEAAQLAPPIYPGGEAIVFMSDFPIAHGTNGLSVPCITPOEFVTHV 461
DB 402 WSLPNTSGRTGHNHVLAPAVAPTFPGEQLLFRSTMPGCGYPNMNLDCLLPQEWLHFY 461
QY 462 NEQAPTRGEAALHYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTQ--TLPINGVFV 519
DB 462 QEAAPASDVALLRFVNPDTGTVLFECKLHKSQYITV-----AHTGPDVLLPNGVFRD 517
QY 520 SVWSRYQLKPVGTAGPACRLGIRR 545
DB 518 SWNVQFYTLAPMGNG-----TGRRA 538

RESULT 3
POLN_RHDV STANDARD; PRT; 2344 AA.
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OX NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome.";
RL Virology 184:664-676(1991).
CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```



QY 479 PDT 481  
DB 637 PDS 639

## RESULT 5

POLN\_MANCV STANDARD; PRT: 2208 AA.  
AC 069014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein].  
OS Manchester virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Gaul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses.";  
RL Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC (RNA)(N).  
CC -!- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X86560; CAA60262.1;  
CC MEROPS; C24 UPW;  
CC InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Endoptase\_C24.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR PRINTS; PR00916; 2CENDOPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? 2208 COAT PROTEIN.  
FT NE\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 10.0%; Score 289; DB 1; Length 2208;  
Best Local Similarity 22.9%; Pred. No. 9.4e-14;  
Matches 138; Conservative 82; Mismatches 24; Indels 138; Gaps 25;  
QY 14 GTSGAGQLVPEANTAEPISEPVAGAA-----TAAATAGQVNMIDPWIMNMYVQAPQGEFT 69

Db 1670 GTTG-----PTTSHVVVANPEQPNGAQRLELAVATGA-----IQSNVPEAIRNCEFA 1716  
QY 70 I-----SPNNTPGDILFDLQGLPHLNPLSLAQAQNYGVGMKVKVLLAGNAFTAGKI 123  
Db 1717 VFTFAWNRDMPGTGTLGSLISLHPNINPYTSLHSGMWAGGSGFEVRLSTSGSGVFAGRI 1776  
QY 124 IISCIPIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFHNDNA-PTMRLV 182  
Db 1777 IASVIPPG--VDPSSIRDPGVLPFAFVDARITEPVSFMPIDVRADVDRHMDGAEPGCSLG 1834  
QY 183 CMLYTFLRASGSSGTPDFVIAGRVLTCPSPDFSEFLFVPP--NVBQKTKPFSVPNPLN 240  
Db 1835 FWYQPLLNPSTAVSTCWVS--VETKGGDFDCLLRPPGQOMENGVSPEGILPRLG 1892  
QY 241 TLSNSRVPSLIKSMVSRDHGQWQV-FQNGRVTLGQLOGTTPTSASQLCKIRGSVFHAN 299  
Db 1893 YSRGNRVGGLVGMILVAEHLKQNRHNSVTFGWSAPVNPMAA-----EIVTNOAHST 1948  
QY 300 GNGY-----NLTEL-DGSPYHAFESPAPI-----GFPDLGECDDHWEASPTTQ 342  
Db 1949 SRHAWLSIGAQNKGPLFGIPNHFDPSCASTVVGAMDTSLGGRPSTGVC-----GPAIS 2002  
QY 343 F-NTGDVIKQINVKQESAFAPH-----LGTIQAQDGLSDSV------N 378  
Db 2003 FQNGGDVYE--NDTPSVMPATYDPLTSGTGVALTNSINPASILALVRISNNDFTSGFAND 2060  
QY 379 TNNIAKLGWSPVSDGH-RGDVDPMVIPRPGYSTLTAEQAQLA-----PPIYPP 424  
Db 2061 KVVVQMSWEMTYGNIQIRGQVTPMSGNTYTFSTGANTLVLMQERMLSYDGHQAIIYSS 2120  
QY 425 GFGEAIVFPMSPDPIAHGTNGLSVFCPTIQEEVTHFVNEQAPTRGEAALLHYLDPDTHRN 484  
Db 2121 QLERTAEVFONDI-----VNIPENSMAYF-----NVETNSA 2151  
QY 485 LGEFLYLPGEFTVCPNNSGCTGQPLPENGVPFVSVWSRFYQLKPVGTA--GPACRLGI 542  
Db 2152 SFQIGIRPDGMYV-----TGGSGIVNVVPLEPETRF-----QYVGILPLSAALSGPSGNGR 2202  
QY 543 RR 544  
Db 2203 AR 2204  
RESULT 6  
COAT\_FCVF9  
ID COAT\_FCVF9 STANDARD; PRT: 671 AA.  
AC P27406;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Coat protein (Capsid protein).  
GN CYP76.  
OS Feline calicivirus (strain F9) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OX NCBI\_TaxID=11981;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92410623; PubMed=1529544;  
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,  
RA Turner P.C.;  
RT "The complete nucleotide sequence of a feline calicivirus.";  
RL Virology 190:443-448(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92339673; PubMed=1633955;  
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
RT "Cloning and sequence determination of the feline calicivirus strain  
RT F9.";  
RL Biochem. Soc. Trans. 20:26S-26S(1992).  
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M86379; AAA79327.1; -  
 DR EM3L; Z11536; CAA77636.1; -  
 DR PIR; B43382; VCVWF9.  
 DR PIR; C45538; C45538.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5E5 CRC64;

Query Match 9.5%; Score 276; DB 1; Length 671;  
 Best Local Similarity 26.4%; Pred. No. 1.8e-13;  
 Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGAATAAATAGOVNMDPWNNYVQAPQGEFTI- 70

Db 127 DGSITA-----PEQGTWGGVIAEPSAQMSTADMATCKSVDSM-----EAFFSFH 173

QY 71 -----SPNPTPGDILDLQGLPHLNFSLHQAOMYNGVNMKVLLAGNAFTAGKIII 125

Db 174 TSVNWTSETQGIKFLKOSGLPLNLYLHAKLYVAMSGSIEVRSISGSGVFGKLA 233

QY 126 SCIPPGF-AAQNTSIAQATMFPHVIAADVRLVLEPIEDVRNVLPHNDNAPTRMLVCM 184

Db 234 IVVPPGDPVQSTSMLOQ---YPHVLFDAQVPEVIFCLPDLRLTLHLSMDTDTTSLIVM 290

QY 185 LY-----TPLRASGSSGCTDFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSPVNPPLN 240

Db 291 VYNDLINPYANDANSSG---IVT--VETKPGDFKFLHLLKPPG-----SMLTHG 335

QY 241 TLNSRVPSLIKSMVSRHQGVQFNGRVTLDGOLQGTPTTSASOLCKIRGSVFHANG 300

Db 336 SIPSDLIPKTSLSLWGNRYNSDITDF-----LTELDSGYHA-----PESPA-PTGFPDL---GEC--- 331

QY 301 GNGYN-----LTELDSGYHA-----PESPA-PTGFPDL---GEC--- 331

Db 373 HFDNQETAGWSTFRFRPISVTITEONGAKLGIGVATDYIVPGIPDGWPDITIPGELIPA 432

QY 332 -DNHM-----EASPTTQFNTGDVIK 350

Db 433 GDVAITNGTNDITATGYDTADIK 458

RESULT 7  
 COAT\_FCVF4  
 ID COAT\_FCVF4 STANDARD; PRT; 668 AA.

AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11980;  
 RN [1]  
 RP Sequence from N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 RA Miyamura K., Yamazaki S., Mikami T.;  
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";

RL Virology 183:810-814(1991).

CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; F90357; BAA14371.1; -

DR PIR; B40481; VCVWFC.

DR InterPro; IPR004005; Calici\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 668 AA; 73589 MW; 85BBD8CB85804E503 CRC64;

Query Match 9.5%; Score 274; DB 1; Length 668;  
 Best Local Similarity 31.0%; Pred. No. 2.6e-13;  
 Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGAATAAATAGOVNMDPWNNYVQAPQGEFTI--T---SPNVT 75

Db 133 PEQGTAVGGVIAEPSAQMSTADMATCKSVDSM-----EAFSEFTSVNWTSET 183

QY 76 PGDILFDLQGLPHLNFSLHQAOMYNGVNMKVLLAGNAFTAGKIIICIPPGF-AA 134

Db 184 QGKILFKQLGSLPLNLYLHLSKLYVAMSGSIEVRSISGSGVFGKLAIVVPGVDV 243

QY 135 QNTISQAQATMFPHVIAADVRLVLEPIEDVRNVLPHNDNAPTRMLVCMXY----TPLR 190

Db 244 QSTSMLOQ---YPHVLFDAQVPEVITFDLRLTLHLSMDTDTTSLIVMNDLINPYA 300

QY 191 AGSSGSGTDFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSPVNPPLNTLSNRVPS 249

Db 301 NDSNSSG---IVT--VETKPGDFKFLHLLKPPG-----SVLTGHSIPS 339

RESULT 8  
 COAT\_FCV6  
 ID COAT\_FCV6 STANDARD; PRT; 668 AA.

AC P27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coat protein (capsid protein).  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11979;  
 RN [1]  
 RP Sequence from N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus";  
 RL J. Virol. 65:5440-5447(1991).  
 RN [2]  
 RP Sequence from N.A.  
 RA Neill J.D.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -





CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL: D00625; BAA00516.1; ALT\_SEQ.  
CC PIR: A34032; GNNY2W.  
CC HSP: P03299; IPOV.  
CC MEROPS: C03.001; -.  
CC MEROPS: C03.020; -.  
CC InterPro: IPR000199; Cys-protease-3C.  
CC InterPro: IPR003138; Pico\_P1A.  
CC InterPro: IPR000081; Pico\_P2A.  
CC InterPro: IPR002527; Pico\_P2B.  
CC InterPro: IPR000605; RNA\_helicase.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR001676; Rnv.  
CC Pfam: PF00073; rhv; 3.  
CC Pfam: PF00548; Cys-protease-3C; 1.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC Pfam: PF00947; Pico\_P2A; 1.  
CC Pfam: PF01552; Pico\_P2B; 1.  
CC Pfam: PF02226; Pico\_P1A; 1.  
CC ProDom: PD001125; Cys-protease-3C; 1.  
CC ProDom: PD001274; Cys\_P2B; 1.  
CC ProDom: PD001306; Pico\_P2A; 1.  
KW Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.  
FT CHAIN 2 69  
FT CHAIN 70 340  
FT CHAIN 341 578  
FT CHAIN 579 879  
FT CHAIN 880 1028  
FT CHAIN 1029 1125  
FT CHAIN 1126 1454  
FT CHAIN 1455 1541  
FT CHAIN 1542 1563  
FT CHAIN 1564 1746  
FT CHAIN 1747 2205  
FT LIPID 2 2  
FT ACT\_SITE 1710 1710  
FT ACT\_SITE 1724 1724  
SQ SEQUENCE 2205 AA; 245701 MW; 2A42AB039E0254AD CRC64;  
  
Query Match 5.1%; Score 146.5; DB 1; Length 2205;  
Best local Similarity 19.2%; Pred. No. 0.0073;  
Matches 109; Conservative 71; Mismatches 188; Indels 199; Gaps 26;  
  
QY 22 VPANTAEPISMEPVAGATAAATAGQVNMIDPWIMNNVVOAQPGFTTSPN-----NTP 76  
DB 365 IPEDVTTPID---IPGEVRNMELAEIDTMTPLNTSQRKNTMDMYRVELNDAAHSDDTP 421  
  
QY 77 GDLFDLQPLHNPFLSH-----LAQYNGWVGNMKVLLAGNAFTAGKIISCIPPG 131  
DB 422 ---ILGLSLSPASDPRLAHTMLGEILNYTHWAGSLKFTFLFCGSMWATKLLVSYAPPG 478  
  
QY 132 FAAQNTSIAQATPFHVIADRVLEPVEPVEDVRNVLFHN--NDNAP-----TWR 180  
DB 479 AKAPE-SRREAMLGTHVIGDGLQSSCTMVVPWISNTYRQTINDSFTEGGYISMFYQTR 537  
  
QY 181 LVMCLYTPLRASGSSSGTDPFVIAGRLVTCPSDFSLFL-----VPPN----- 224  
DB 538 VVYPLSTPRKMD-----ILGFVSAC--NDFSRLRLRDTTHISQEVMPQGLGLDIE 585

QY 225 --VEQTKPSPFVNPLNPLNTLSNSR-----VPS-----LIKSMVSRDHQ 262  
DB 586 GNVVEGVTNRALTPLTPVNNPLDTRSSGPAHSKETPALTAVETGATNPLVPSDFVQTRH-- 643  
QY 263 MVQ-----FQNGRVTLDDGQOGTTPT-SASOLCKIRGSVF----- 296  
DB 644 V:OKRSTSESTVESFFARGACVAILIENDAPTRRASKLSFVWKIITYKDTVQLRRKLEFF 703  
QY 297 -----HANGGYNLTDLGSPYHAFESPAPIGFPDLG---ECDWHMEA 337  
DB 704 TYSRFDEMEFTFVVTSTNYTDANNGHALNQVQIMY-----IPPGAPLPGRNDYTWQTS 757  
QY 338 SPITQENTGVIRKQINVKQESAFAPHLGTQA-----DGLSDSVSVNTNMIKLGWSPVS 392  
DB 758 NPSVFTYTGAPPARISV-----PYVGIANAYSHFYDGFQKV----- 793  
QY 393 DCHRGDVPWVPIPRYGSTLTLEAAQLAPPIYPGFGFGEAIVFEMSFDEPIAHCTNGLSVPCTI 452  
DB 794 -----PLAQASTEGDSL-----YGAA-----GSLAV----- 820  
QY 453 PQEFVTHFVNEQAPTRGEAALLHYLDP 479  
DB 821 -----RVVNDHNPTKLTSKIRVYMKP 841  
  
RESULT 11  
POLG\_POL2L  
ID POLG\_POL2L STANDARD; PRT; 2207 AA.  
AC P06210;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C, P3A; Genome-linked protein VP6; Picornain 3C  
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D  
DE (EC 2.7.7.48)].  
OS Poliovirus type 2 (strain Lansing).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=12084;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86115399; PubMed=3003384;  
RA la Monica N., Meriam C., Racanelli V.R.;  
RT "Mapping of sequences required for mouse neurovirulence of poliovirus  
RT type 2 Lansing";  
RL J. Virol. 57:515-525(1986).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the  
CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
CC (RNA)(N).  
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M12197; AAA46912.1; -.  
CC PIR: A29507; GNNY5P.  
CC HSP: P03299; IPOV.  
CC MEROPS: C03.001; -.  
CC MEROPS: C03.020; -.  
CC

DR InterPro: IPR000199; Cys-protease-3C.  
DR InterPro: IPR003138; Pico\_P1A.  
DR InterPro: IPR000081; Pico\_P2A.  
DR InterPro: IPR002527; Pico\_P2B.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR001676; Rhv.  
DR Pfam: PF00073; rhv; 3.  
DR Pfam: PF00548; Cys-protease-3C; 1.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
DR Pfam: PF00947; Pico\_P2A; 1.  
DR Pfam: PF01552; Pico\_P2B; 1.  
DR Pfam: PF02226; Pico\_P1A; 1.  
DR ProDom: PD001125; Cys-protease-3C; 1.  
DR ProDom: PD001274; Pico\_P2B; 1.  
DR ProDom: PD001306; Pico\_P2A; 1.  
KW polyprotein; coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.  
FT CHAIN 2 69  
FT CHAIN 70 340  
FT CHAIN 341 578  
FT CHAIN 579 879  
FT CHAIN 880 1028  
FT CHAIN 1029 1125  
FT CHAIN 1126 1454  
FT CHAIN 1455 1541  
FT CHAIN 1542 1563  
FT CHAIN 1564 1746  
FT CHAIN 1747 2207  
FT LIPID 2 2  
FT ACT\_SITE 1710 1710  
FT ACT\_SITE 1724 1724  
SQ SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;

Query Match 5.08; Score 145.5; DB 1; Length 2207;  
Best Local Similarity 18.8; Pred. No. 0.0087;  
Matches 106; Conservative 72; Mismatches 193; Indels 193; Gaps 24;

QY 22 VPEANTAPISMEPVAGATAAATAGQVNMIDPMNMNYVQAPOGETTSPN-----NTP 76  
DB 365 IPFDVTPPID--IPGEVRNMELAEIDTMIPLNLTNRKNTDMYRVELNDAAHSDTP 421  
QY 77 GDILFDLQHLNPLFLSH-----LAQYNGWGNMKVLLAGNATAGKILICPPG 131  
DB 422 ---ILCLSLSPASDRLAHTMLGEITLNYTHWAGSLKFTFLFCGSMATCKLLVSYAPP 478  
QY 132 FAAQNISTAQATMPHVIAADVRLIEVPLEDVRLVFNH--NDNAP-----TMR 180  
DB 479 AEAPK-SRKEAMLGTHVLDIGLQSSCTMVVPWISNTTYROTINDSFTEGYISMFYQTR 537  
QY 181 LVCMLYTLRASGSSGTDPEVIAGRVLTCPSPDFSELF-----LVP 222  
DB 538 VVVPLSTPKMD-----ILGFVSAC--NDFSVRLRDTTHISOEAMPOGLGLDIE 585  
QY 223 PNVEQKTPFSPVNPPLNTLSNR-----VPS-----LIKSMVYSRDHQ 262  
DB 586 GVVEGVTRNALTPLTPANLPDPTOSSGPAHSEKTPALTAVETGATNPVPSDVTQTRH-- 643  
QY 263 MVQ-----FQGRVTLGLOGTTPTS-ASOLCKIRSVF----- 296  
DB 644 VIQKTRSESTVESFARGACVAIEVDNDAPTFRKSLFWSKWTITVDVOLRRKLEFF 703  
QY 297 -----HANGNGYNLTLDGSPYHAFESPAPIGFPDGLGEDWHEASPT 340  
DB 704 TVSRFDEMEFTFVTSNYTDANNGHALNQYQIMYPGAPIPGKWDY---TWQTSSNPS 760  
QY 341 TQNTGDKVQKINVKQESAFPHLGTIOA-----DGLSDSVNTNMIAKLGNWSPVSDGH 395  
DB 761 VFYTYGAPARLSV-----PYVGIANAYSHFYDGFVKV----- 793  
QY 396 RGDVDPWIPRYGSLTEAALAPDIYPPGCEAIVFMSDFPIAHGTLGSLVPCITPQE 455

DB 794 -----PLAQOASTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820  
QY 456 FVTHFVNEQAPTRGEAALLHYLDP 479  
DB 821 ---RVVNDHNPTKLTSKIRVYMKP 841

RESULT 12  
POLG\_POL3L  
ID POLG\_POL3L STANDARD; PRT: 2206 AA.  
AC P03302; O84783; O84784; O84785; O84786; O84787; O84788; O84789;  
AC O84790; Q98592; Q98593; Q98594;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; P3A; Genome-linked protein VP3; Picornain 3C  
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D  
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D  
DE (EC 2.7.7.48)].  
OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=12088;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P3/Leon/37;  
RX MEDLINE=84170338; PubMed=6324200;  
RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,  
RA Schild G.C., Almond J.W.;  
RT "Comparison of the complete nucleotide sequences of the genomes of  
RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin  
RT vaccine derivative P3/Leon 12a1b.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P3/Leon 12A[1]B;  
RX MEDLINE=83299339; PubMed=6310508;  
RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,  
RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;  
RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison  
RT with poliovirus type 1.";  
RL Nucleic Acids Res. 11:5629-5643(1983).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.  
RX MEDLINE=95120467; PubMed=7820548;  
RA Grant R.A., Hiramath C.N., Filman D.J., Syed R., Andries K.,  
RA Hogle J.M.;  
RT "Structures of poliovirus complexes with anti-viral drugs:  
RT implications for viral stability and drug design.";  
RL Curr. Biol. 4:784-797(1994).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.  
RA Hiramath C.N., Grant R.A., Filman D.J., Hogle J.M.;  
RT "Binding of the antiviral drug win51711 to the Sabin strain of type-3  
RT poliovirus - structural comparison with drug-binding in rhinovirus-  
RT 14.";  
RL Acta Crystallogr. D 51:473-489(1995).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the  
CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -!- PTM: SPECIFIC ENYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS  
CC SHOWN.  
CC -!- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE  
CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.



```
DR InterPro: IPR003138; Pico_P1A.
DR InterPro: IPR000081; Pico_P2A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001676; Rnv.
DR Pfam: PF00073; rhv; 3.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR ProDom: PD001125; Cys-protease-3C; 1.
DR ProDom: PD001274; Pico_P2B; 1.
DR ProDom: PD001306; Pico_P2A; 1.
DR PolyProtein: Coat protein; Core protein; Hydrolase; Thiol protease; Myristate.
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69
FT CHAIN 70 319
FT CHAIN 320 561
FT CHAIN 562 871
FT CHAIN 872 1014
FT CHAIN 1015 1113
FT CHAIN 1114 1443
FT CHAIN 1444 1532
FT CHAIN 1533 1554
FT CHAIN 1555 1737
FT CHAIN 1738 2194
FT LIPID 2
FT ACT_SITE 1701 1701
FT ACT_SITE 1715 1715
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE0673C CRC64;

Query Match 5.0%; Score 143.5; DB 1; Length 2194;
Best Local Similarity 19.4%; Pred. No. 0.012;
Matches 94; Conservative 57; Mismatches 166; Indels 167; Gaps 17;

QY 70 ISPNTPGDILP-----DIQL-GPHLNPFLSHLAQWNGWGMKVKVLGAGNATAGKII 124
DB 390 ISAQMDQLLIPNLDIQLEPLNTLLGNISRYTHWSGLEMTMFCGSEMTTKGLI 449
QY 125 ISCIPEGAAQNIATAQTMFPHVIADVRVLEPIEVLPLEYRNVLFNHNDNAPMTMLVCM 184
DB 450 ICYTPGGSSPTDRM-QAMLAHVVMDEGLQSSITIIIPWISSGSHYR-----M 496
QY 185 LYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFVSNPLPLNTLSN 244
DB 497 FNTDAKAINAVG-----YVTC-----FMQTNL----- 519
QY 245 SRVPSLIKSMVSRDHGMQVQFNGRVTLTGLOGLCTPTTSASQLCKIRGVSFHANGNGY 304
DB 520 -----VAPGGAADQCYIVGMV---AAKKDF 541
QY 305 NLTELDGSPYHAFESPAPIGPDLGCDWHEASPTTQFNIGDVIKQINVKQESAFAPHL 364
DB 542 NULMRDS-----PDIGQSAILPEQAATQI--GEIVKTVANTVSEIKAEI 586
QY 365 GTIQADGLSDSVNTNMIKLGWSPVSDGHRGDVDPWIPRYGSTLTLEAAQLAPPIYPP 424
DB 587 GVIPS-----LNNAVETGATSNTEP-----EAAIQTRTVINMH 618
QY 425 GGEIAIV-----FPMSPDPIAHNGTNGLSVPCCTIPEQVTFHVN--EQAPTRGGAAL 473
DB 619 GTAECLEVNFLGRSALVCMRSFEYKNHSTSTS---SIQKNFFIWTNLTRELVLRRKMEI 675
QY 474 LHYLDPTDHRNLGEPKLYPEGMTCVPSNSGTGPQTLPINGVFEVSVWSRYFLKPVGT 533
DB 676 FTYLFEDT-----EITIVPTLRLFSSNSVFSGLPNLTLMQAMV-----PTGA 718
QY 534 AGPA 537
DB 719 RKPS 722
```

```
RESULT 14
POLG_POL32 STANDARD; PRT; 2206 AA.
AC P06209;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VP6; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strain 23127).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87010550; PubMed=3020156;
RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W.,
RA Stanway G.;
RT "The nucleotide sequence of a type 3 poliovirus isolated during a
RT recent outbreak of poliomyelitis in Finland.";
RL J. Gen. Virol. 67:2093-2102(1986).
CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensed@isb-sib.ch).
CC
CC EMBL; X04468; CAA28155.1; -
CC PIR; A27245; GNNY27.
CC HSSP; Q84730; LPVC.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC
CC InterPro: IPR000199; Cys-protease-3C.
CC InterPro: IPR003138; Pico_P1A.
CC InterPro: IPR000081; Pico_P2A.
CC InterPro: IPR002527; Pico_P2B.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR001676; Rnv.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 579 937
```

```
FT CHAIN 579 878 COAT PROTEIN VP1.
FT CHAIN 879 1027 CORE PROTEIN P2A.
FT CHAIN 1028 1124 CORE PROTEIN P2B.
FT CHAIN 1125 1453 CORE PROTEIN P2C.
FT CHAIN 1454 1540 CORE PROTEIN P3A.
FT CHAIN 1541 1562 CORE PROTEIN P3B.
FT CHAIN 1563 1745 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1746 2206 PICORNAIN 3C.
FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;

Query Match 4.8%; Score 139.5; DB 1; Length 2206;
Best Local Similarity 20.8%; Pred. No. 0.025;
Matches 96; Conservative 64; Mismatches 182; Indels 119; Gaps 21;

QY 22 VPENTAEPTSMPEVAGATAAATAGOVNMIDPWIMNMYVOAQOGFTI---SPNTPGD 78
DB 365 IPEDVTPPID---IPGEVKNVMELEIDTMIPLNLENTKRTMDMYRVLSDSANLSGP 421
QY 79 ILFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVLLAGNAFTAGKIIISCIPPGFA 133
DB 422 ILC-LSUSPAADPRLSHMTLGEVLYNTHWAGSLKFTFLFCGSMATGKLLVAYAPG-A 479
QY 124 AQNISIAQATMFPHVIADVRVLEPIEDVRNVLPFH--NNDNAPTMRLVCMLY----- 186
DB 480 QPPTSREKMLGTHVINDLGLSSCTMVVPWISNVYTRQTTQDSFTGEGYISMFYQTRIV 539
QY 187 TPLRASGSSGTDDEVIAGRVLTCPSPDFSLFL-----VPPNVE-----OKT 229
DB 540 VPL-----STPKAMDMGLGFVSAC--NDFSRLLRDTHISQAAMPQGVDDLITEVAQNA 591
QY 230 KPESVP---NLP-----LNTLS-----NSRVPS 249
DB 592 LALSCLPQSNLPDTPKASGPAHSKEVPTLTAETGATNPLVPSDTVOTRIVIOQRSSSES 651
QY 250 LIKSM-----MVSRDHGO---MVG--FQNGRVT-----LDGQLQGTTPTSASOLCK 290
DB 652 TIESFFARGACVAIEVDNEQATNVQKLFATWRTITYKDVQLRRKLEFFTSYRFDNEFT 711
QY 291 IRGSVFHANGNGVNLTELDGSPYHAFESAPGFPDGLGECNDHMEASPTQRTNGTVIK 350
DB 712 FVVTANETNSNGHALNOVQIIMYPGAPTCKSWDDY---TWOTSSNPSIFTYGAAPA 768
QY 351 QINVKQESAPAFHGLTIOA-----DGLSDVSVNTNMLAKLG 386
DB 769 RISV-----PYVGLANAYSHFDGFAKVPKASDANDQVG 802

RESULT 15
POLG_POLIS STANDARD; PRT: 2209 AA.
AC P03301; Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;
AC Q84888; Q84889; Q84890;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.18) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 1 (strain Sabin).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299876; PubMed=6310545;
RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,
RA Genba Y., Nakano Y., Imura N.;
RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1
RT strain genome.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM
CC THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION
CC PROCESS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01150; CAA24465.1; -
CC PIR; A03899; GNNY3P.
CC MEROPS; C03.001; -
CC MEROPS; C03.020; -
CC
CC InterPro; IPR000199; Cys-protease-3C.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR000081; Pico_P2A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR01205; RNA_pol_P3D.
CC InterPro; IPR001676; Rhv.
CC
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC
CC Polyprotein; Coat protein; Core protein; Transferrase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 341 COAT PROTEIN VP3.
FT CHAIN 342 579 COAT PROTEIN VP3.
FT CHAIN 580 881 COAT PROTEIN VP1.
FT CHAIN 882 1030 COAT PROTEIN P2A.
FT CHAIN 1031 1127 CORE PROTEIN P2B.
FT CHAIN 1128 1456 CORE PROTEIN P2C.
FT CHAIN 1457 1543 CORE PROTEIN P3A.
FT CHAIN 1544 1565 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1566 1747 PICORNAIN 3C.
FT CHAIN 1748 2209 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE.
FT ACT_SITE 1712 1712 PROTEASE (POTENTIAL).
FT ACT_SITE 1726 1726 PROTEASE (POTENTIAL).
SQ SEQUENCE 2209 AA; 246576 MW; 9EC1EF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;
Best Local Similarity 19.3%; Pred. No. 0.036;
Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPENTAEPTSMPEVAGATAAATAGOVNMIDPWIMNMYVOAQOGFTI---GDI 79
DB 366 LPEFDVTPPID---IPGEVKNMELAEIDTMIPIFDLSAKKNTMEMYRVLSDKPTDDP 422
QY 80 LFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVLLAGNAFTAGKIIISCIPPGFAA 134
```

Db	423	ILCLUSLSPASDPLRSLHMLGEILNYTHWAGSLUKFTFLFCGSMATGKLLVSYAPPG- AD 481
Qy	135	QNISIAQATPHPHVIADRVLEPIEVPLEDVRNVLPHNN-DNAPT-----MRLVC 183
Db	482	PPKKRKEAMLGTHVIWDJIGLSSCTWVPWPIWNTTYRQIDDSFTBGGVISVFYQTRIVV 541
Qy	184	MLTYPLRASGSSGTDFFVIAGRVLTCPSDFSF-LFLVPPNVEQK----- 228
Db	542	PLSTPREMD-----ILGFVSAC--NDFSRLMRDTHIEQAKLAAGLGQMLESMI 589
Qy	229	-----TKPFSVPNLPLNLTLSNR-VPS-----LTKSMVSRDQGQMVQ 265
Db	590	DNTVRETVGAATSRDALPNTPEASGPAHSKEIPALTAVETGATNPLVPSTDVQTRH--VVQ 647
Qy	266	-----FQNGR-----VFLDGLQCTTPTSASQLCKTRGS----- 294
Db	648	HRSRSESSIESFARGACVAIIIVDNSASTKNKDKLFTVWKITYKDTVOLRRKLFFIYS 707
Qy	295	-----VPHAN---GGNGYNLTLDGGPYHAFESPAPIGFPDLGECDDHMEASPTTQF 343
Db	708	RDFMEFTFVVTANFTETNNGHALNQVYIMYVPPGAPVPEKWDY---TWQTSSNPSIFY 764
Qy	344	NTGDVIKIQINVKQESAFAPHLGTIOA-----DCLSDVSYNTNTIAKLGWVSPVSDGHRGD 398
Db	765	TYGTAPARISV-----PYVGISNAYSHFYDGFSGV-----PLKD----- 798
Qy	399	VDPWVIPRYGSTLTEAAQLAPPYPPGFGEAIVFFMSDFEPIAHGTNGLSVPCCTIQEFVFT 458
Db	799	-----QSAALGDSLY-----GAA---SLNDFGI-----LAV 821
Qy	459	HFVNEQAPTRGEAALLHYLDP 479
Db	822	RVVNDHNPTKVTSKIRVYLPK 842

Search completed: March 10, 2003, 18:26:52  
Job time : 26 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:17:40 ; Search time 43 Seconds  
(without alignments)  
1688.874 Million cell updates/sec

Title: us-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPNTMDTSGAGQ.....YQLKPVCCTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2896	100.0	545 22 AAB49700	Small round struct
2	1979	68.3	544 22 AAB49703	Small round struct
3	1977.5	68.3	546 22 AAB49702	Small round struct
4	1950	67.3	530 15 AAR50972	Norwalk virus stra
5	1938	66.9	530 22 AAB49701	Small round struct
6	1900	65.6	530 15 AAR57091	Small round virus
7	1225	42.3	540 22 AAB49706	Small round struct
8	1178.5	40.7	548 22 AAB49705	Small round struct
9	1174.5	40.6	535 22 AAB49707	Small round struct
10	1172.5	40.5	542 22 AAB49708	Small round struct

11	1157.5	40.0	539	22	AAB49704	Small round struct
12	1157	40.0	548	23	AAU91272	Norwalk virus asso
13	1154.5	39.9	550	22	AAB49709	Small round struct
14	1013	35.0	541	22	AAU91270	Small round struct
15	321	11.1	579	18	AAU91271	RHDV capsid protei
16	276	9.5	547	22	AAU91273	Feline calicivirus
17	276	9.5	671	22	AAU91274	Feline calicivirus
18	272.5	9.4	669	22	AAU91275	Amino acid sequenc
19	272	9.4	623	22	AAU91276	Feline calicivirus
20	272	9.4	668	22	AAU91277	Amino acid sequenc
21	269	9.3	623	22	AAU91278	Feline calicivirus
22	269	9.3	668	12	AAU91279	Feline calicivirus
23	269	9.3	668	22	AAU91280	Feline calicivirus
24	262.5	9.1	622	22	AAU91281	Feline calicivirus
25	154	5.3	2206	13	AAU91282	True type 3 poliov
26	147	5.1	40	23	AAU91283	Norwalk virus asso
27	136	4.7	40	23	AAU91284	Norwalk virus asso
28	132.5	4.6	2209	3	AAU91285	Sequence encoding b
29	125.5	4.3	2179	7	AAU91286	Sequence encoding b
30	116.5	4.0	1707	22	AAU91287	Novel human diagno
31	112.5	3.9	2164	9	AAU91288	Peptides translate
32	112.5	3.9	2164	9	AAU91289	Sequence of the vi
33	110	3.8	1037	22	AAU91290	Drosophila melanog
34	110	3.8	1234	22	AAU91291	Drosophila melanog
35	110	3.8	2016	22	AAU91292	Drosophila melanog
36	107.5	3.7	2599	21	AAU91293	Neisseria meningit
37	106	3.7	1302	20	AAU91294	Human MEK1 protei
38	106	3.7	1302	20	AAU91295	Human MEK1 protei
39	106	3.7	1493	20	AAU91296	Murine MEK1 prote
40	106	3.7	1493	20	AAU91297	Murine MEK1 prote
41	106	3.7	1593	23	AAU91298	Murine MEK1-2 M
42	105.5	3.6	239	9	AAU91299	Sequence of rhinov
43	103	3.6	1559	21	AAU91300	Arabidopsis thalia
44	103	3.6	1596	21	AAU91301	Arabidopsis thalia
45	103	3.6	1752	21	AAU91302	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAB49700  
ID AAB49700 standard; protein; 545 AA.  
AC AAB49700;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 1.  
EE Small round structured virus; SRSV; food poisoning.  
FF Small round structured virus.  
OS Small round structured virus.  
XX  
XX WO2000079280-A1.  
XX  
XX 28-DEC-2000.  
XX  
XX 22-JUN-2000; 2000WO-JP04095.  
XX  
XX 22-JUN-1999; 99JP-0175928.  
XX  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI: 2001-080848/09.  
XX N-PSDB; AAF29141.  
XX  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
XX strains for investigation of food poisoning outbreaks, contains  
XX antibodies -

```

XX PS Claim 1; Page 40-42; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 545 AA;

Query Match 100.0%; Score 2896; DB 22; Length 545;
Best Local Similarity 100.0%; Pred. No. 3, le-250;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDAPTNNDGTSAGQLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNYY 60
DB 1 MMASKDAPTNNDGTSAGQLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNYY 60
QY 61 VOAPQGEFTISPNNTPGDIPLDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGDIPLDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSEVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSEVPLNPLN 240
QY 241 TILSNRVPFLIKSMVSRDHQWQFONGRVTLTGQGLQGTTPTSASQLCKIRGSVFHANG 300
DB 241 TILSNRVPFLIKSMVSRDHQWQFONGRVTLTGQGLQGTTPTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTFNTGDIKQINVKQESAF 360
DB 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTFNTGDIKQINVKQESAF 360
QY 361 APHLGTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQLAPP 420
DB 361 APHLGTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQLAPP 420
QY 421 IYPPGGEAIVFMSDFPIAHGTNGLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 480
DB 421 IYPPGGEAIVFMSDFPIAHGTNGLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 480
QY 481 THRNLEGEFLKXPEGMTCVPNSSGTGQTLPIINGVFVSVWSRFFYQLKPVGTAGPACRL 540
DB 481 THRNLEGEFLKXPEGMTCVPNSSGTGQTLPIINGVFVSVWSRFFYQLKPVGTAGPACRL 540
QY 541 GTRRS 545
DB 541 GTRRS 545

RESULT 2
AAB49703
ID AAB49703 standard; protein; 544 AA.
XX AC AAB49703;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 4.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO200079280-A1.

```

```

XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP04095.
XX PR 22-JUN-1999; 99JP-0175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WI: 2001-080848/09.
XX DR N-PSDB: AAF29144.
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies -
XX PS Claim 1; Page 47-49; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 544 AA;

Query Match 68.3%; Score 1979; DB 22; Length 544;
Best Local Similarity 66.8%; Pred. No. 3, le-168;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPTNNDGTSAGQLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNYY 60
DB 1 MMASKDAPTNNDGTSAGQLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNYY 60
QY 61 VOAPQGEFTISPNNTPGDIPLDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGDIPLDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSEVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSEVPLNPLN 240
QY 241 TILSNRVPFLIKSMVSRDHQWQFONGRVTLTGQGLQGTTPTSASQLCKIRGSVFHANG 300
DB 241 TILSNRVPFLIKSMVSRDHQWQFONGRVTLTGQGLQGTTPTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTFNTGDIKQINVKQES 358
DB 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTFNTGDIKQINVKQES 358
QY 358 QRVNLNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTFNTGDIKQINVKQES 358
QY 359 AFAPHLGTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQLA 418
DB 359 AFAPHLGTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQLA 418
QY 419 PLYPPGGEAIVFMSDFPIAHGTNGLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 474
DB 419 PLYPPGGEAIVFMSDFPIAHGTNGLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 474
QY 475 HYLDPDTHRNLEGEFLKXPEGMTCVPNSSGTGQTLPIINGVFVSVWSRFFYQLKPVGTAG 534
DB 475 HYLDPDTHRNLEGEFLKXPEGMTCVPNSSGTGQTLPIINGVFVSVWSRFFYQLKPVGTAG 534
QY 535 GPA-CRLGIRR 544
DB 535 GPA-CRLGIRR 544

```



[illegible]

QY 61 VQAPQGEFTISPNNTPGDILDLQGLHLPFLSHLAQMYNGVGNMKVVKVLLAGNAFTA 120  
 DB 61 VQAPQGEFTISPNNTPGDILDLQGLHLPFLSHLAQMYNGVGNMKVVKVLLAGNAFTA 120  
 QY 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVNRVLFHNND-NAPTM 179  
 DB 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVNRVLFHNND-NAPTM 179  
 QY 180 RLVCMLYTLPASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNPPL 239  
 DB 180 RLVCMLYTLPASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNPPL 239  
 QY 240 NTLNSRVPSLIKSMVSRDHQVQFNGRVTLDGQLOGTTPTSASQICKIRGSVFHAN 299  
 DB 240 NTLNSRVPSLIKSMVSRDHQVQFNGRVTLDGQLOGTTPTSASQICKIRGSVFHAN 299  
 QY 239 SSLNSRAPLPISMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 DB 239 SSLNSRAPLPISMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 QY 300 GGNGYNLTDELSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESA 359  
 DB 300 GGNGYNLTDELSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESA 359  
 QY 295 NGTVINLTDELDTGTFPHFEGPAPIGFDPDLGCGCDWHIN---MTQFCHSSQOTQYDVTTPDT 351  
 DB 295 NGTVINLTDELDTGTFPHFEGPAPIGFDPDLGCGCDWHIN---MTQFCHSSQOTQYDVTTPDT 351  
 QY 360 FAPHLGTQADGLSDVSVNTNMIAKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQAP 419  
 DB 360 FAPHLGTQADGLSDVSVNTNMIAKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQAP 419  
 QY 352 FVPHLGSIQANGIG---SGNYGVLSWSPSPSHPSQVLDLKKIPNYGSSITEATHLAP 407  
 DB 352 FVPHLGSIQANGIG---SGNYGVLSWSPSPSHPSQVLDLKKIPNYGSSITEATHLAP 407  
 QY 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
 DB 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
 QY 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVYISHLASEQAPTVGEAALLHYVDP 464  
 DB 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVYISHLASEQAPTVGEAALLHYVDP 464  
 QY 480 DTHRNLTGEFKLYPBGFMTCVNSSGTGPTPLPINGVFWVSWRSRYOLKPVGTAGPA-C 538  
 DB 480 DTHRNLTGEFKLYPBGFMTCVNSSGTGPTPLPINGVFWVSWRSRYOLKPVGTAGPA-C 538  
 QY 465 DTGRNLTGEFKAYPDGFLTCVNGASSGQQPLPINGVFWVSWRSRYOLKPVGTASSARG 524  
 DB 465 DTGRNLTGEFKAYPDGFLTCVNGASSGQQPLPINGVFWVSWRSRYOLKPVGTASSARG 524  
 QY 539 RLGIIR 544  
 DB 539 RLGIIR 544  
 QY 525 RLGLRR 530  
 DB 525 RLGLRR 530

RESULT 5  
 AAB49701 standard; protein; 530 AA.  
 ID AAB49701  
 AC AAB49701;  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 2.  
 KW Small round structured virus; SRSV; food poisoning.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 PD 28-DEC-2000.  
 PF 22-JUN-2000; 2000WO-JP04095.  
 PR 22-JUN-1999; 99JP-0175928.  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 FA (DENK-) DENKA SEIKEN KK.  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 DR WPI; 2001-080848/09.  
 XX N-PSDB; AAF29142.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies  
 XX Claim 1; Page 42-45; 84pp; Japanese.  
 PS  
 XX

CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks.  
 XX Sequence 530 AA;  
 QY Query Match 66.9%; Score 1938; DB 22; Length 530;  
 QY Best Local Similarity 67.0%; Pred. No. 1.6e-164;  
 QY Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;  
 QY 1 MMASKDAPNTMDGTSCAGOLVPEANTAPISMEPVAGAAATAAGOVNIDPWIMNNY 60  
 DB 1 MMASKDAPNTMDGTSCAGOLVPEANTAPISMEPVAGAAATAAGOVNIDPWIMNNY 60  
 QY 61 VQAPQGEFTISPNNTPGDILDLQGLHLPFLSHLAQMYNGVGNMKVVKVLLAGNAFTA 120  
 DB 61 VQAPQGEFTISPNNTPGDILDLQGLHLPFLSHLAQMYNGVGNMKVVKVLLAGNAFTA 120  
 QY 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVNRVLFHNND-NAPTM 179  
 DB 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVNRVLFHNND-NAPTM 179  
 QY 180 RLVCMLYTLPASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNPPL 239  
 DB 180 RLVCMLYTLPASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNPPL 239  
 QY 240 NTLNSRVPSLIKSMVSRDHQVQFNGRVTLDGQLOGTTPTSASQICKIRGSVFHAN 299  
 DB 240 NTLNSRVPSLIKSMVSRDHQVQFNGRVTLDGQLOGTTPTSASQICKIRGSVFHAN 299  
 QY 239 SSLNSRAPLPISMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 DB 239 SSLNSRAPLPISMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 QY 300 GGNGYNLTDELSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESA 359  
 DB 300 GGNGYNLTDELSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESA 359  
 QY 295 NGTVINLTDELDTGTFPHFEGPAPIGFDPDLGCGCDWHIN---MTQFCHSSQOTQYDVTTPDT 351  
 DB 295 NGTVINLTDELDTGTFPHFEGPAPIGFDPDLGCGCDWHIN---MTQFCHSSQOTQYDVTTPDT 351  
 QY 360 FAPHLGTQADGLSDVSVNTNMIAKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQAP 419  
 DB 360 FAPHLGTQADGLSDVSVNTNMIAKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQAP 419  
 QY 352 FVPHLGSIQANGIG---SGNYGVLSWSPSPSHPSQVLDLKKIPNYGSSITEATHLAP 407  
 DB 352 FVPHLGSIQANGIG---SGNYGVLSWSPSPSHPSQVLDLKKIPNYGSSITEATHLAP 407  
 QY 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
 DB 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
 QY 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVYISHLASEQAPTVGEAALLHYVDP 464  
 DB 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVYISHLASEQAPTVGEAALLHYVDP 464  
 QY 480 DTHRNLTGEFKLYPBGFMTCVNSSGTGPTPLPINGVFWVSWRSRYOLKPVGTAGPA-C 538  
 DB 480 DTHRNLTGEFKLYPBGFMTCVNSSGTGPTPLPINGVFWVSWRSRYOLKPVGTAGPA-C 538  
 QY 465 DTGRNLTGEFKAYPDGFLTCVNGASSGQQPLPINGVFWVSWRSRYOLKPVGTASSARG 524  
 DB 465 DTGRNLTGEFKAYPDGFLTCVNGASSGQQPLPINGVFWVSWRSRYOLKPVGTASSARG 524  
 QY 539 RLGIIR 544  
 DB 539 RLGIIR 544  
 QY 525 RLGLRR 530  
 DB 525 RLGLRR 530  
 RESULT 6  
 AAR57091 standard; protein; 530 AA.  
 ID AAR57091  
 AC AAR57091;  
 DT 05-OCT-1994 (first entry)  
 DE Small round virus SRSV/KY/89 capsid protein.  
 KW pathogen; acute gastroenteritis; food poisoning;  
 KW seafood contamination; diagnostic assay; human calicivirus;  
 KW small round virus; SRSV; KY89; Norwalk virus; capsid protein.  
 XX Small round virus SRSV/KY/89.  
 XX WO9405700-A.

PD 17-MAR-1994.  
 XX 07-SEP-1993; 93WO-US08447.  
 XX 07-SEP-1992; 92US-0941365.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX Estes MK, Graham DY, Jiang X, Matson DO;  
 PI WPI; 1994-101125/12.  
 DR N-PSDB; AAQ56832.  
 XX DNA from Norwalk and related viruses - used for preparing prods.  
 PT for use in diagnostic assays, detection and vaccines for Norwalk  
 PT and related viruses  
 XX  
 PS Example 7; Fig 13a; 156pp; English.  
 XX  
 CC The known sequence for Norwalk virus was used to obtain the  
 CC sequence of other Norwalk-related viruses such as SRSV/KY/89, an  
 CC agent from a stool from an outbreak of gastroenteritis in Japan  
 CC in 1989. The 2516 nucleotide cDNA sequence includes part of the  
 CC polymerase region and the capsid region of the genome; the deduced  
 CC amino acid sequences are AAR57092 and AAR57091, respectively. Expression  
 CC of fragments and derivs. of Norwalk-related viruses permits  
 CC development of diagnostic assays to detect antibodies, antigens,  
 CC viral genetic material or antivirals.  
 XX  
 SQ Sequence 530 AA;  
 Query Match 65.6%; Score 1900; DB 15; Length 530;  
 Best Local Similarity 65.7%; Pred. No. 4e-161;  
 Matches 362; Conservative 66; Mismatches 95; Indels 28; Gaps 9;  
 QY - MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60  
 Db 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60  
 QY 61 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120  
 Db 61 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120  
 QY 121 GKIIISCPGPPGFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLPHND-NAPTM 179  
 Db 121 GKIIISCPGPPGFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLPHND-NAPTM 179  
 QY 180 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 239  
 Db 180 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 239  
 QY 240 NTLNSRVPSPSLIKSMWVRDHQWVQFONGRVTLDGLOGTTTTSASQLCKIRGSVFHAN 299  
 Db 240 NTLNSRVPSPSLIKSMWVRDHQWVQFONGRVTLDGLOGTTTTSASQLCKIRGSVFHAN 299  
 QY 300 GNGYNLTDELGSPYHAFESPAPIGPFDPLGECDDHMEAE-----SPTQFNTGDVIRKQIN 354  
 Db 300 GNGYNLTDELGSPYHAFESPAPIGPFDPLGECDDHMEAE-----SPTQFNTGDVIRKQIN 354  
 QY 353 KQESAFAPHLICTIOADGLSDVSVTNMIKLGWSPVSDGHRGDVDPWIPRYGSTLFEA 414  
 Db 353 KQESAFAPHLICTIOADGLSDVSVTNMIKLGWSPVSDGHRGDVDPWIPRYGSTLFEA 414  
 QY 402 -----PHLGSIQANGIG-----SGNTIGLWSVSPSPHSGQVOLKWPNGSITAE 402  
 Db 402 -----PHLGSIQANGIG-----SGNTIGLWSVSPSPHSGQVOLKWPNGSITAE 402  
 QY 474 AQLAPPIYPPGFEAIVFMSDFPIAHGTNGLSVPCITPOEFYTHEVNEQAPTGEAALL 474  
 Db 474 AQLAPPIYPPGFEAIVFMSDFPIAHGTNGLSVPCITPOEFYTHEVNEQAPTGEAALL 474  
 QY 403 THLASVSPGFEGLVFPFMSKIP---GPGGDSLPCLLPGLGYISHLASEQAPTVEGEP 459  
 Db 403 THLASVSPGFEGLVFPFMSKIP---GPGGDSLPCLLPGLGYISHLASEQAPTVEGEP 459  
 QY 475 HYLDDPDTNRNLGEFLYPEGFWTCVPSNGSGTGPOTLPINGVFVSVMSREFYOLKPVGTA 534  
 Db 475 HYLDDPDTNRNLGEFLYPEGFWTCVPSNGSGTGPOTLPINGVFVSVMSREFYOLKPVGTA 534  
 QY 460 HVDVDPDTNRNLGEFLYPEGFWTCVPSNGSGTGPOTLPINGVFVSVMSREFYOLKPVGTA 519  
 Db 460 HVDVDPDTNRNLGEFLYPEGFWTCVPSNGSGTGPOTLPINGVFVSVMSREFYOLKPVGTA 519  
 QY 535 GPA-CRLGIR 544  
 Db 535 GPA-CRLGIR 544

Db 520 STARGRLGLRR 530  
 RESULT 7  
 AAB49706  
 ID AAB49706 standard; protein; 540 AA.  
 XX AAB49706;  
 AC AAB49706;  
 XX  
 XX 04-APR-2001 (first entry)  
 DT  
 XX Small round structured virus protein SEQ ID 7.  
 DE  
 XX Small round structured virus; SRSV; food poisoning.  
 KW  
 XX Small round structured virus.  
 OS  
 XX WO200079280-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX 22-JUN-2000; 2000WO-JP04095.  
 PF  
 XX 22-JUN-1999; 95JP-0175928.  
 PR  
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 PI WPI: 2001-080848/09.  
 DR N-PSDB; AAF29147.  
 XX  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies -  
 XX  
 XX Claim 1; Page 54-57; 84pp; Japanese.  
 PS  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks.  
 XX  
 SQ Sequence 540 AA;  
 Query Match 42.3%; Score 1225; DB 22; Length 540;  
 Best Local Similarity 46.7%; Pred. No. 1e-100;  
 Matches 261; Conservative 89; Mismatches 173; Indels 35; Gaps 14;  
 QY 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60  
 Db 1 MMASKADPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60  
 QY 61 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120  
 Db 61 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120  
 QY 57 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 116  
 Db 57 VQAPGEGTISPNNTPGDILDLGLPHLPNLPFLSHLAQMYNGVGNMKVLLAGNAFTA 116  
 QY 121 GKIIISCPGPPGFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLPH-NNDNAPTM 179  
 Db 121 GKIIISCPGPPGFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLPH-NNDNAPTM 179  
 QY 180 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 239  
 Db 180 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 239  
 QY 177 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 234  
 Db 177 RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFVLPVNPVEOKTKPFSVPL 234  
 QY 240 NTLNSRVPSPSLIKSMWVRDHQWVQFONGRVTLDGLOGTTTTSASQLCKIRGSVFHAN 299  
 Db 240 NTLNSRVPSPSLIKSMWVRDHQWVQFONGRVTLDGLOGTTTTSASQLCKIRGSVFHAN 299  
 QY 235 GELNSRFPDLSDMVTSPNESIVVQPNQGRVTLDELGLLGTTLQACNICSIRKVTGV 294  
 Db 235 GELNSRFPDLSDMVTSPNESIVVQPNQGRVTLDELGLLGTTLQACNICSIRKVTGV 294



XX	Small round structured virus protein SEQ ID 9.
DE	
XX	
KW	Small round structured virus: SRSV: food poisoning.

RESULT 11  
 AAB49704  
 ID AAB49704 standard; protein; 539 AA.  
 XX  
 AC AAB49704;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Small round structured virus protein SEQ ID 5.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP04095.  
 XX  
 PR 22-JUN-1999; 99JP-0175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX  
 PI (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 DR WPI; 2001-080848/09.  
 XX  
 DR N-PSDB; AAF29145.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies -  
 XX  
 PS Claim 1; Page 50-52; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks.  
 XX  
 SQ Sequence 539 AA;  
 Query Match 40.0%; Score 1157.5; DB 22; Length 539;  
 Best Local Similarity 45.6%; Pred. No. 1.1e-94;  
 Matches 253; Conservative 83; Mismatches 192; Indels 27; Gaps 14;  
 QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNYY 60  
 Db 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNYY 60  
 QY 61 VOAPCGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGMVKVLLAGNAFTA 120  
 Db 57 VOAPCGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGMVKVLLAGNAFTA 116  
 QY 121 GKIIISCIPPGFAAGNTSIAQATMFHVIADVLEPIEVLDPEDVNRNLFH-NNDNAPTM 179  
 Db 117 GKIIISCAVPPNPPTEGSLSPQVTHFIIIVDVRQLEPVLIPLDVRRNFIHYNQSNDSIT 176  
 QY 180 RLVCMXYTLPRASSSSGTDFFVIAGRVLCPSDFSLFLVPPNVQKTRPFSVPLNPL 239  
 Db 177 KLIAMLYTLPLRA--NNAGDDVFTVSCVLRPSPDFDFLFPVPTVESRTPFTVPLTV 234  
 QY 240 NTLNSRVPSLTKMWSRDRHGOMVQFQNGRVTLTGLOGGTTPTSASOLCKTRGSVRHAN 299  
 Db 235 EEMSNSRPIPLEKLYTCPSAFVVOQNGRCRTDGLVLTQLSALNLCIFRGDYTHIA 294  
 QY 300 GNGY--NLTELDSGPYHAFES-PAPIGFPD-LGECWHEMASPTQFNTGDTVIKQINVK 355

Db 295 GSHDYTMNLASQNSNYDPTTEIPAPLGTDFVGKIQGLM--TQTTREDGSGTRAHKATVS 352  
 QY 356 QESA-FAPHLGTIOADGLSDVSVNTNMIKLGWSPVSDG--HRCGDVDPWIPRYGSTLT 412  
 Db 353 TGSVHFTPKLGSVOYTTDTNDFOTGQNTKFTPVGVQIDGNHONHPQOQWLPNYSGRGT 412  
 QY 413 EAAQLAPPITYPPGFEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNQAPTRGEAA 472  
 Db 413 HNVHLAPAVAPTFEGEQLLFFRSTMPGSCGYPMNMLDCLLPQEWVQHFQCEAAQAQSDVA 472  
 QY 473 LLHYLDPTHNLGFEKLYPGEFWTCVPNSSGTGCPOTL--PINGVVFVSVWSRYQLKP 530  
 Db 473 LLRFVNDPTGVLFECKLHKSGYTV----AHTGPHDLVIPPNGYFRFDSWVNOFYTLAP 528  
 QY 531 VGTAGPACRLGIRRS 545  
 Db 529 MNGA-----GRRRA 538  
 RESULT 12  
 AAU91272  
 ID AAU91272 standard; Protein; 548 AA.  
 XX  
 AC AAU91272;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Norwalk virus associated polynucleotide #1.  
 XX  
 KW Norwalk virus; monoclonal antibody; geno group I; geno group II;  
 KW immunological detection; food; viral infection.  
 XX  
 OS Norwalk virus.  
 XX  
 PN JP2002020399-A.  
 XX  
 PD 23-JAN-2002.  
 XX  
 PF 10-JUL-2000; 2000JP-0208151.  
 XX  
 PR 10-JUL-2000; 2000JP-0208151.  
 XX  
 PA (OSAP) OSAPA PREFECTURE.  
 PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.  
 PA (IATR) IATR LAB INC.  
 XX  
 DR WPI; 2002-287412/33.  
 XX  
 PT A monoclonal antibody useful in the immunological detection and  
 PT diagnosis of Norwalk virus infection -  
 XX  
 PS Disclosure; Page 12-13; 24pp; Japanese.  
 CC The invention describes a monoclonal antibody recognising Norwalk virus,  
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the  
 CC capsid protein molecule of geno group I and geno group II. The antibody  
 CC is useful for immunological detection and quantitative analysis of  
 CC Norwalk virus in foods and the serum of infected patients. This sequence  
 CC represents a Norwalk virus associated protein described in the invention.  
 XX  
 SQ Sequence 548 AA;  
 Query Match 40.0%; Score 1157; DB 23; Length 548;  
 Best Local Similarity 45.7%; Pred. No. 1.3e-94;  
 Matches 254; Conservative 83; Mismatches 177; Indels 42; Gaps 15;  
 QY 2 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNYY 61  
 Db 3 MASSRAAFSN-DGAAG--LVPEINN-EAALDPVAGAAIAAPLTGQNIIDPWIMNMFV 57  
 QY 62 QAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGMVKVLLAGNAFTA 121

Db 58 QAPGGEFTVSPRNSGCVGLLNLELGPENIPYLAHLARMYNGYAGFEVQVVLACNAFTAG 117  
 QY 122 KILISCIPTPGFAAQTISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPT-MR 180  
 Db 118 KILFAAIPPNFIDNLSAAQITMCPHVIDVRVLEPIEVPLEDVRNVLPHNDNAPT-MR 177  
 QY 181 LVCMLYTPLRASGSSCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNPLN 240  
 Db 178 LIAMLYTPLRASGSSCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNPLN 235  
 QY 241 TLSNSRVPSLTKSMVSRDRGQWQFQNGRVLTDLGLOGTTPTSASOLCKIRGSVFEHANG 300  
 Db 236 EMSNSRFPVLESHTSPENIVVOQNGRVLTDLGLOGTTPTSASOLCKIRGSVFEHANG 295  
 QY 301 G-----NGY---NLTELDGSPYHAFES-PAPIGPDL-GSCDWHMEASPTQ 342  
 Db 296 RASDAQADIATPRFNYYWVQVLDNLGTPYPAEDIPGLGTPDRGKV-----FGVASQ 350  
 QY 343 FNTGDIKQINVKQESA---FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDG--HRG 397  
 Db 351 RNPDSSTRAHEAKVDTTGTAGRTPKLGSLEISTESS-DEQONQPTRE---TPVGIGVDNEA 406  
 QY 398 DVDPRVPRYGSTLTAAQAAPPYPPGFGAIVFPFMSDFPIAHGTNGLSVPTIPQEFV 457  
 Db 407 DFQMSLUPDYSGQTHNNLAPAVAPNFTGEOLLEFRSOLPSSGGRSNGVLDLVPQEW 466  
 QY 458 THEVNEQAPTREGAALLHYLDPDTHRLNGEFLKYPEGFMTCVPSNGSGTGPOTLPIGVFV 517  
 Db 467 QHYQESAPACTQVALRVYVNPDTGRVLFELAKHLKLGEMTIAKN--GDSPTVPPNGYER 524  
 QY 518 FVSWSRFVYQKPKVGT 533  
 Db 525 FESWVNPFTYIAPMGT 540

RESULT 13  
 AAB49709 standard; protein; 550 AA.  
 XX  
 AC AAB49709;  
 XX  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 10.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP04095.  
 XX  
 PR 22-JUN-1999; 99JP-0175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 DR WP1; 2001-080848/09.  
 DR N-PSDB; AAF29150.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies  
 XX  
 PS Claim 1; Page 62-64; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAB20141 - AAB20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks.

XX  
 SQ Sequence 550 AA;  
 Query Match 39.9%; Score 1154.5; DB 22; Length 550;  
 Best Local Similarity 45.6%; Pred. No. 2.2e-94;  
 Matches 259; Conservative 73; Mismatches 175; Indels 61; Gaps 14;

QY 1 MMASKO-APINMDGTSAGAGOLYPEANTAPFISMEPVAGAAATAAGTACOVNMDPMTNN 59  
 Db 1 MKMASNDAAASN-----DGAANLYPEAND-EVMALEPVVGASIAAPVQGNLIIDWIWREN 55  
 QY 60 YVQAPQGEFTVSPRNSGCVGLLNLELGPENIPYLAHLARMYNGYAGFEVQVVLACNAFT 119  
 Db 56 FVQAPQGEFTVSPRNSGCVGLLNLELGPENIPYLAHLARMYNGYAGFEVQVVLACNAFT 115  
 QY 120 AGKIIISCIPTPGFAAQTISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLPHNDNAPT 178  
 Db 116 AGKIIIFAAVPPHPPVENISAAQITMCPHVIDVRVLEPIEVPLEDVRNVLPHNDNAPT 175  
 QY 179 MRLVCMYTPLRASGSSCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNPLN 238  
 Db 176 MRLVAMLYTPLRASGSSCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNPLN 232  
 QY 239 LNTLSNSRVPSLTKSMVSRDRGQWQFQNGRVLTDLGLOGTTPTSASOLCKIRGSVFEHANG 297  
 Db 233 LGEALSNSRFPVLESHTSPENIVVOQNGRVLTDLGLOGTTPTSASOLCKIRGSVFEHANG 292  
 QY 298 -----ANGNGYNTLDELGSPYHAFES-PAPIGPDLGSCDWHMEASPTQ 342  
 Db 293 TARAADSTDSQARARNHPLHVQVKNLDGTOYDPTDIPAVLGAIQF-----KGTV 342  
 QY 343 FNTGDIKQINVKQESA---FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDG--HRG 385  
 Db 343 FGVASQ-RDVSQEQEQGHYATRAHEAHIDTDPKYAPKLGTLIKSGSD-DEFTNQPIRF 400  
 QY 386 GWSPVSDGHRGVDVPHVPIRYGSTLTAAQAAPPYPPGFGAIVFPFMSDFPIAHGTNG 445  
 Db 401 ---TPVGMDG-NNRWQLPFDYSGRLTLNMLAPAVSPSPGGRILUFFRSIYVPSAGGYS 456  
 QY 446 LSVPTIPQEEVTHEVNEQAPTREGAALLHYLDPDTHRLNGEFLKYPEGFMTCVPSNGSGT 505  
 Db 457 GYIDCLIPQEWVQHYQESAPACTQVALRVYVNPDTGRVLFELAKHLKLGEMTIAKN--GDSPTVPPNGYER 517  
 QY 506 GPCTLPINGVFEVSWSRFVYQKPKVGT 533  
 Db 515 NPLVVPNGYERFEAWGNQFYTLAPMGS 542

RESULT 14  
 AAB49710 standard; protein; 541 AA.  
 XX  
 AC AAB49710;  
 XX  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 11.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP04095.  
 XX

PR 22-JUN-1999; 99JP-0175928.  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 DI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI; 2001-080848/09.  
 DR N-PSDB; AAF29151.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies  
 XX  
 PS Claim 1; Page 64-66; 84pp; Japanese.  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks.  
 XX  
 SQ Sequence 541 AA;  
 Query Match 35.0%; Score 1013; DB 22; Length 541;  
 Best Local Similarity 41.4%; Pred. No. 9.9e-82;  
 Matches 243; Conservative 81; Mismatches 155; Indels 108; Gaps 19;  
 QY 1 MKMASDAPNTMDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNY 60  
 DB 1 MKMASDAPNTMDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNY 60  
 QY 61 VOAPQGETISPNTPGDTILFLOLPHLNPFLSHLAQWNGWGNKYKVLNAGNFTA 120  
 DB 57 VOAPAGEFTVSPRNSPGEILLDLGPDNLPLAHLAMYNHAGGGEVOIVLAGNFTA 116  
 QY 121 KIIISCIPGFAQAQNIATQATMFPFHVADRVLEPIEPLVDNRVLFH-NNDNAPTM 179  
 DB 117 KIIIFAAPGFPYENLSPSTMTGPHVILIDVROGLFLLPMPDIWNFFHNOGNDPKL 176  
 QY 180 RLVCMLYTLPRASGSSGTDPEVIAGRVLTCPSPDFGLFVLPNVBOKTKPESVNLPL 239  
 DB 177 RLVAELYTLPLA--NNSGDDVFTVSCRVLTKPSDFEFLVPTVESKTKQFALPLKI 234  
 QY 240 NTLNSRVPSLIKSMVRDHQMTQFONGRVTLDCQLQGTTPTSASQLCKIRGVSFHAN 299  
 DB 235 SEMTNSRFPVPVDMYATARNQVQVQNGRVTLDCGLLGTTPLLAVNICKFKRGEVIAKN 294  
 QY 300 GG-NGY---NLTELDGSPYHAFE-SPAPIGFPDGLGCDWHM-----EASPTQ--- 342  
 DB 295 GDVRSYRMDMEITNCTPTDPTDTPGIGSPDFQGLFCVASQRKNKNEQNPATRAHEA 354  
 QY 343 -FNTG--DVTKOINVKQESAFAPHL-----GTQADGLSDVSVNTNMIK 384  
 DB 355 IINTGGDLCPQISSEIYLTSPILKNTCPQLPQSLRGRTILI--RSDNGHCHDMVG- 411  
 QY 385 LGWVSPV-----SDGHRGDVDPWIPRYGSLTLPAAQLAPPIYPGFG 426  
 DB 412 ---TSPTTTPWQQRRCRSGSCSGHRRVPV-PVVMNR----- 447  
 QY 427 GEALVFFMSPDPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPTNRNIG 486  
 DB 448 -----VTIVLSHKSFGSTSKLPQ-----LNLNRP-----LRIENEDTGRVLF 488  
 QY 487 EFKLYPGCFMTCVPNSSGTPQTLTINGVFVSVSWSVSEFFQLKPVGT 533  
 DB 489 EARLHKGQFTIVA--HTGDNPIVMPNPGYFPEAWVNOFYSLAPVGT 533  
 RESULT 15  
 AAW08143

ID AAW08143 standard; Protein; 579 AA.  
 XX  
 AC AAW08143;  
 XX  
 DT 12-MAR-1997 (first entry)  
 XX  
 DE RHDV capsid protein.  
 XX  
 KW Canarypox virus; vaccinia virus; ALVAC; NVVAC; vector; attenuation;  
 KW calicivirus; capsid; antigen; RHDV; vaccine; diagnosis; therapy.  
 XX  
 OS Rabbit haemorrhagic disease virus strain Saone et Loire.  
 XX  
 PH Key Location/Qualifiers  
 FT Modified-site 45  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 281  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 308  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 369  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 393  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 430  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 474  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 481  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 FT Modified-site 502  
 FT /label= Glycosylation  
 FT /note= "putative N-glycosylation site"  
 XX WO639177-A1.  
 XX  
 PD 12-DEC-1996.  
 XX  
 PF 03-JUN-1996; 96WO-IB00721.  
 XX  
 PR 06-JUN-1995; 95US-0471025.  
 XX  
 PA (VIRO-) VIROGENETICS CORP.  
 XX  
 PI Fischer L, Legros F, Paoletti E;  
 XX  
 DR WPI; 1997-042857/04.  
 DR N-PSDB; AAT42749.  
 XX  
 PT Modified recombinant virus contg. exogenous calicivirus DNA - useful  
 PT in vaccines and for in vitro prodn. of calicivirus antigens, for  
 PT generation of therapeutic or diagnostic antibodies  
 XX  
 PS Example 12; Fig 12a; 150pp; English.  
 XX  
 CC The rabbit haemorrhagic disease virus (RHDV) capsid protein  
 CC (AAW08143) amino acid sequence was deduced from the putative capsid  
 CC gene (AAT42749) obt'd. from RHDV strain Saone et Loire. The capsid  
 CC protein can be expressed from novel recombinant, attenuated  
 CC vaccinia virus and canarypox virus vectors, esp. based on NVVAC and  
 CC ALVAC, useful as vaccines of improved safety. The protein may  
 CC also be expressed in vitro and used to generate therapeutic or  
 CC diagnostic antibodies.  
 XX  
 SQ Sequence 579 AA;



```

Query Match      11.1%; Score 321; DB 18; Length 579;
Best Local Similarity 28.3%; Pred No. 9.2e-20;
Matches 132; Conservative 61; Mismatches 171; Indels 102; Gaps 21;

QY 12 MDGTSGAGQLVPENATAPIS-----MEPVAGAAATA-----ATAG----- 47
Db 1 MEGKARAPOGEAAGTATTASVGTGTTDGMDFGVVATTSVITAENSSASIAATAGIGGPPQ 60
QY 48 QVNMIDPWIMNNYVQAPQGEFTTISPNTPCDILFDLQGLPHLNPFLSHLAQMYNGWYGNM 107
Db 61 QVDQQTWRTNFTY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGM 117
QY 108 KVKVLLAGNAFTAGKIIICIPGFAAQNISTA---QATMPPHVIADVRLPIEVPLED 164
Db 118 QRFIVAGSGVFGRLVAAPVPG-----IEIGPLEVRFQFPHVVIDARSLEPVITIMPD 172
QY 165 VRNVLFH-NNDNAPTMLVCMXYPL--RASGSSGTDPPVIAGRVLTCPSPDFSEFLV 221
Db 173 LRPNYHPTGDPGLVPTLVSVYNNLINPFGGSTS-----AIQVTVETTPSEDPEFVMIR 227
QY 222 PPNVEQTKPFSVNLPLNTLSNSRVPSLIKSMVSRDHCOMVQFQ-----NGRVT 272
Db 228 APS--SKTVDSISAGLLTT-----PVLTVGNDNRWNGQIVGLQPVPGGFSTCNRHN 279
QY 273 LDGQLQGTTPTSASQLCKIRGSVFHANGNGYNLTLEL-----DGSPIHAFESP-----APIG 324
Db 280 LNGSTYGWSSPRFADIDHRRGSASY-SGNATNVLFQFWYANAGS---AIDNPISQVAPDG 335
QY 325 FPDLGECDDHMEASPTTQFNTGDVIOINVKQESAFAPHLGTIOADGLSDVSNTNMIK 384
Db 336 FPDMSFVPFNGPGIPAGWVGFGALWNSN-----SGAPNVTVQA-----YE 377
QY 385 LGWVSPVSDHGRGDVDPWVPRYGSILTEAAQ-LAPPIYPPGGEA 429
Db 378 LGFAT---GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 412

```

Search completed: March 10, 2003, 18:26:16  
Job time : 45 secs

